University of Alberta

Identification of new targets of the regulatory protein BldD in *Streptomyces coelicolor*

by

Claire Marielle Galibois

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Microbiology and Biotechnology

Department of Biological Sciences

Edmonton, Alberta

Spring, 2007

Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file   Votre référence
ISBN: 978-0-494-29676-9
Our file   Notre référence
ISBN: 978-0-494-29676-9

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Abstract

Streptomyces are a group of organisms renowned for their unusual growth, erecting sporulating aerial hyphae, and for production of most of the naturally occurring antibiotics. In *S. coelicolor*, both processes are controlled by the *bld* (for bald) genes. Mutants of *bldD* are unable to produce either aerial hyphae or antibiotics; this gene encodes a DNA-binding regulator known to repress transcription of its target genes in early development and to regulate the spatial expression of at least one.

New targets of BldD regulation have been found through microarray analysis following *bldD* induction in the *bldD* mutant strain 1169. The analysis identified *prsI* which encodes a putative anti-sigma factor and *arsI* which encodes a putative anti-sigma factor antagonist. Both *prsI* and *arsI* were repressed in early development, and expression of *prsI* was restricted to the aerial structures at all times tested. Mutants of *arsI* exhibited a *bld* phenotype, and deletion of *prsI* was presumed lethal. DNase I footprinting was used to confirm the BldD binding site upstream of *arsI*. Reanalysis of the regions protected by BldD in previously known targets combined with the binding sites in new targets has allowed the BldD binding consensus to be refined to AGTGA (n_m) g/tCACg/c, where n is any base in a sequence of variable length (m).

BldD was found to activate the transcription of *sigQ*, which encodes an extracytoplasmic function sigma factor. *sigQ* is the first gene shown to be activated, rather than repressed, by BldD. There are numerous putative BldD

binding sites in the *sigQ* promoter, although the exact sequences bound have not been determined. Binding of BldD to *sigQ* was demonstrated both *in vivo* and *in vitro.* BldD binding to two of the newly identified BldD targets, *prsI* and *arsI*, was significantly weaker than BldD binding to previously published targets, although *sigQ* was bound strongly. Other genes identified as potential BldD targets from the microarray analysis are tabulated and briefly discussed.

In this work, new BldD targets have been identified and characterized, and potential roles for *prsI* and *arsI* in regulating development are proposed.

Acknowledgements

I would first like to thank my supervisor, Dr. Brenda Leskiw, for teaching, encouraging and guiding me through my time in her lab. She has been mentor, teacher and friend, and without her support and encouragement I would not have made it through to the end of this Ph.D. Her advice for me in dealing with students was "be kind," and I cannot think of anyone who better lives her words.

Thank you also to the members of the Leskiw lab present when I started - Marie Elliot, Julie Stoehr, and Dawn Bignell. I felt welcomed into a family, and patience with endless questions was certainly appreciated! The family grew over my time in the lab, and I have been fortunate to have worked with such wonderful people - graduate students Kimberley Colvin, Kent Gislason, Linda Bui, Archana Parashar and Wisdom Ayidzoe, undergraduate students Janiffer Song, Leon Lau, Chris Fung and Curtis Hrdlicka and technicians Annie Wong, Troy Locke and Marie Kaplan. In particular I would like to thank Kim for her supportive friendship, especially in the craziness of combining research time with preparing for a wedding, Archana for friendship and sharing her cultural experience, and Linda for many long discussions.

Many people have helped with my research. Thanks to Troy Locke for help with BldD purification and help with miscellaneous questions. Kent Gislason continued the BldD purification, and did the lion's share of the EGFP analysis begun by Chris Fung, besides helping with the recalcitrant *prsI* overexpression and deletion projects. Linda Bui generated all of the knockouts described in this thesis, besides continuing the BldD purification and optimizing ChIP for *S.*

coelicolor, I used her technique in my experiments. Finally, the collaboration of Dr. Camilla Kao and Nitsara Karoonuthaisiri at Stanford University was invaluable in the microarray analysis.

I would also like to thank Anna Szenthe, co-ordinator of the lab portion of Microbiology 265 and my supervisor as a teaching assistant for most of my degree. I learned a great deal about teaching during this time, and I enjoyed our discussions and appreciated the rides to early morning labs.

Finally, I would like to thank my family and friends for support, for laughter, and for letting me describe at length the fascinating world of *Streptomyces* genetics. To my parents: you provided the foundation on which I have built my life, and your unconditional love has helped me through the tough times. To my husband, Rob: you have been with me through the best and worst times of my life (both during grad school!) and have between them helped me to remember the joys of life and of discovery. Thank you.

Table of Contents

Page

Chapter 1. Introduction **2**

1.1 General overview **2**

1.2 The *whi* genes: regulators of sporulation **6**

1.3 Regulation of antibiotic production **14**

1.4 The *bld* genes: regulators of morphological differentiation and antibiotic production **18**

1.5 The *ram* and *chp* genes: surfactants and surface proteins **29**

1.6 Thesis objectives **32**

Chapter 2. Materials and Methods **35**

2.1 Bacterial Strains, Plasmids and Growth Conditions **35**
2.1.1 *Streptomyces coelicolor* and *Escherichia coli* strains 35
2.1.2 Plasmids 35
2.1.3 *E. coli* growth conditions 35
2.1.4 *S. coelicolor* growth conditions 42
2.1.5 *S. coelicolor* spore and mycelial stocks 42
2.1.6 Preparation of competent *E. coli* cells 44
2.1.8 Generation of *S. coelicolor* protoplasts 44

2.2 Transfer of DNA **44**
2.2.1 Transformation of competent *E. coli* cells 44
2.2.2 Conjugation into *S. coelicolor* 45
2.2.3 Transformation of *S. coelicolor* protoplasts 46

2.3 DNA manipulations **46**
2.3.1 Isolation of plasmid DNA from *E. coli* 46
2.3.2 Isolation of chromosomal DNA from *S. coelicolor* 46
2.3.3 Agarose gel electrophoresis 47
2.3.4 Purification of DNA from agarose gels 47
2.3.5 Polyacrylamide electrophoresis 48
2.3.6 Purification of DNA from polyacrylamide gels 48
2.3.7 Cloning of DNA fragments 49
2.3.8 Polymerase chain reaction 50
2.3.9 Labelling of DNA 51

2.4 DNA analysis **56**
2.4.1 Southern transfer 56
2.4.2 Southern hybridization 56
2.4.3 Colony hybridization 57
2.4.4 Sequencing 58

2.5 RNA purification and analysis **59**
2.5.1 RNA purification and quantification 59
2.5.2 Northern analysis 61
2.5.3 S1 nuclease mapping 62
2.5.4 Primer extension 65
2.5.5 Microarray analysis 66

2.6 DNA-protein interactions **67**

2.6.1 Electrophoretic Mobility Shift Assay (EMSA) 67
2.6.2 DNase I footprinting 68
2.6.3 Chromatin immunoprecipitation (ChIP) 70

2.7 Gene disruptions 75

2.8 EGFP fluorescence microscopy 80

Chapter 3. Results 83

3.1 Bioinformatic analysis did not reveal new BldD targets 83
3.1.1 Parameters of the search for BldD binding sites in the *S. coelicolor* genome sequence 83
3.1.2 Results of the bioinformatics search 86

3.2 Microarray analysis revealed new BldD targets 90
3.2.1 Construction of the vector for controlled overexpression of *bldD* 90
3.2.2 Insertion of pAU244 into a wild type strain 93
3.2.3 Microarray analysis of wild type *S. coelicolor* in the presence of *bldD* overexpression 93
3.2.4 Insertion of pAU244 or pIJ6902 into the *bldD* point mutant (1169) 97
3.2.5 Microarray analysis of RNA isolated from the *bldD* mutant (1169) *S. coelicolor* after induced *bldD* overexpression 100
3.2.5.1 Confirmation of repression of known targets by *bldD* induction 100
3.2.5.2 Hybridization of microarrays and analysis of results 107

3.3 *prsI* is a target of BldD regulation 116
3.3.1 Sequence analysis of the *prsI* promoter region 116
3.3.2 Analysis of *prsI* transcription patterns 119
3.3.3 Binding of BldD to the *prsI* promoter region 128
3.3.4 DNase I footprinting of the *prsI* promoter region 139

3.4 *arsI* is a target of BldD regulation 145
3.4.1 Sequence analysis of the *arsI* promoter region 145
3.4.2 Analysis of *arsI* transcription patterns 148
3.4.3 Binding of BldD to the *arsI* promoter region 152
3.4.4 Determination of the site of BldD binding in the *arsI* promoter by DNase I footprinting 155

3.5 *sigI* is not a BldD target 159
3.5.1 Sequence analysis of the *sigI* promoter region 159
3.5.2 Analysis of *sigI* transcription patterns 159

3.6 Creation of *prsI*, *arsI*, and *sigI* null mutants 160
3.6.1 PCR-targeted mutagenesis of *prsI* 160
3.6.2 PCR-targeted mutagenesis of *arsI* 169
3.6.3 PCR-targeted mutagenesis of *sigI* 175

3.7 Analysis of cell specificity of transcription 183
3.7.1 Examination of *bldD* expression within the colony 183
3.7.2 Examination of spatial regulation of *prsI* transcription 186
3.7.3 Examination of spatial regulation of *arsI* transcription 192

3.8 *sigQ* as a potential activated BldD target 193
3.8.1 Effect of *bldD* induction on *sigQ* transcription and sequence analysis of the *sigQ* promoter region 193
3.8.2 Determination of the *sigQ* transcription start point 199
3.8.3 Binding of BldD to the *sigQ* promoter region 200
3.8.4 PCR-targeted mutagenesis of *sigQ* 203

3.9 ChIP analysis of putative BldD targets 209
3.9.1 Liquid Culture ChIP 210
3.9.1.1 Experimental setup 210
3.9.1.2 Demonstration of *in vivo* binding by BldD 215
3.9.2 Solid Culture ChIP 218
3.9.2.1 Experimental setup 219
3.9.2.2 Demonstration of *in vivo* binding to known BldD targets 219
3.9.2.3 Examination of *in vivo* BldD binding to promoter regions of negative control genes 222
3.9.2.4 Examination of *in vivo* binding of BldD to *sigQ* 224
3.9.2.5 Examination of *in vivo* BldD binding to *prsI* 227
3.9.2.6 Examination of *in vivo* binding of BldD to *arsI* 228

Chapter 4. Discussion 230

4.1 *prsI/arsI* - New Targets found by microarray analysis 231

4.2 Refined BldD binding consensus sequence 242

4.3 Weak binding of BldD to *prsI* and *arsI* 249

4.4 BldD as an activator 254

4.5 BldD as a dual-function regulator 261

4.6 Conclusion 265

Bibliography 267

Appendix A: Potential BldD targets identified by microarray analysis of RNA isolated from the *S. coelicolor bldD* mutant (1169) after induced *bldD* overexpression 281

A.1 *sco0751* (*scf81.10c*) as a potential BldD target 281

A.2 *sigE* (*sco3356, sce94.07*) as a potential BldD target 289

Appendix B: Putative targets of BldD chosen for PCR-targeted mutagenesis 291

B.1 *sco5718* (*sc3c3.04*) as a potential BldD target 291

B.2 *sco2529* (*scc117.02*) as a potential BldD target 301

Appendix C: *prsI* overexpression attempt 305

Appendix D: BldD overexpression/purification attempts 307

List of Figures

Page

Chapter 1

1.1 *Streptomyces coelicolor* life cycle ... 5

Chapter 2

2.1 ChIP (chromatin immunoprecipitation) procedure, diagram ... 72

2.2 Deletion of genes in the chromosome using PCR-targeted mutagenesis ... 78

Chapter 3

3.1 Chromosomal gene structure and sequence surrounding *bldD* ... 85

3.2 Construction of a plasmid vector (pAU244) for inducible expression of *bldD* ... 92

3.3 Microarray analysis after induced overexpression of *bldD* in wild type *S. coelicolor* ... 95

3.4 The *S. coelicolor bldD* mutant (1169) is complemented by induced expression of *bldD* from plasmid pAU244 ... 99

3.5 Northern analysis of *bldD* transcripts performed on RNA isolated for microarray analysis ... 102

3.6 *bldD* induction represses *bldN* expression in a *bldD* mutant strain ... 104

3.7 High resolution S1 nuclease protection mapping of *bldD* transcripts performed on RNA isolated for microarray analysis (from the *bldD* mutant 1169 with pIJ6902 or pAU244) ... 106

3.8 Microarray analysis of global gene expression changes in response to *bldD* induction in a *bldD* mutant genetic background ... 110

3.9 Microarray results plotted as the $\log_2$-fold change in expression relative to the uninduced sample at each time point post-induction ... 114

3.10 Gene structure in the chromosomal area surrounding *prsI* ... 118

3.11 Sequence of the *prsI* coding region and surrounding area, including part of the *arsI* coding region ... 121

3.12 High resolution S1 nuclease mapping of the *prsI* transcription start points 124
3.13 Primer extension to analyze levels of expression from the *prsIp1* and *prsIp2* promoters in RNA isolated from wild type (916) and *bldD* mutant (1169) *S. coelicolor* strains over a time course of development 127
3.14 EMSA demonstrating BldD binding to its own promoter 130
3.15 Strategy for EMSA of BldD binding to the *prsI* promoter region 133
3.16 EMSA of BldD binding to the *prsI* promoter region near *prsIp1* 135
3.17 EMSA of BldD binding to the *prsI* promoter region 138
3.18 DNase I footprinting of the *bldD* promoter using the 257 bp MAE11-12 probe 141
3.19 Attempt at DNase I footprinting of BldD binding to *prsIp1* using the 142 bp probe CGA43-42 144
3.20 Sequence of the *arsI* coding region and surrounding area 147
3.21 BldD regulation of *arsI* transcription shown by S1 nuclease mapping and primer extension 150
3.22 BldD binds to the *arsI* promoter region 154
3.23 DNase I footprinting of the *arsI* promoter region using a 69 bp probe (CGA47-48) 158
3.24 High resolution S1 nuclease mapping of the *sigI* transcription start point using a 312 bp probe (CGA62-63) 162
3.25 PCR-targeted mutagenesis of four *S. coelicolor* genes 165
3.26 Confirmation of spontaneous spectinomycin resistance in pAU256 exconjugants 168
3.27 Confirmation of replacement of the wild type *arsI* gene with the apramycin resistance cassette in the chromosome of *S. coelicolor* 171
3.28 *arsI*, but not *sigI*, mutants have a bald phenotype 174
3.29 Confirmation of replacement by PCR-targeted mutagenesis of the wild type *sigI* gene in cosmid SCE25 with the apramycin resistance cassette 177

3.30 Confirmation of replacement of the wild type *sigI* gene with the apramycin resistance cassette in the chromosome of *S. coelicolor* 180
3.31 Δ*sigI* mutants show no change in response to osmotic shock 182
3.32 Construction of a *bldD-egfp* transcriptional fusion 185
3.33 Fluorescence microscopy of the wild type *S. coelicolor* strain M145 expressing the *bldD-egfp* transcriptional fusion 188
3.34 Fluorescence microscopy of the wild type *S. coelicolor* strain M145 expressing the *prsI-egfp* transcriptional fusion 191
3.35 Chromosomal gene environment of *sigQ* and promoter sequence showing putative BldD binding sites 195
3.36 Sequence of the *sigQ* coding region and surrounding area 198
3.37 BldD binds to the *sigQ* promoter region 202
3.38 *sigQ* EMSA with CGA40-52 (83 bp) 205
3.39 Confirmation of replacement of the wild type *sigQ* gene with the apramycin resistance cassette in the chromosome of *S. coelicolor* 208
3.40 Agarose gel electrophoresis of sonicated DNA fragments from total cell-free extracts of *S. coelicolor* generated for ChIP experiments 213
3.41 Chromatin immunoprecipitation (ChIP) assay of BldD binding to target DNA in cell extracts from liquid cultures 217
3.42 Chromatin immunoprecipitation (ChIP) assay of BldD binding to target DNA in cell extracts from solid cultures 221
3.43 Chromatin immunoprecipitation (ChIP) assay of BldD binding to target DNA in cell extracts from solid cultures 226

Chapter 4

4.1 BldD binding sites in seven target promoters 244

4.2 Sequence logo illustrating the previously published and refined BldD binding consensus sequences 246

4.3 Sequence logo illustrating the sequences weakly and strongly bound by BldD 253

Appendix

A.1 Microarray results, plotted as the $\log_2$-fold change in expression relative to the uninduced sample at each time point post-induction..... 288

B.1 *sco2529* and *sco5718* are repressed in response to *bldD* induction ... 293

B.2 Confirmation of replacement of the wild type *sco5718/sc3c3.04* gene with the apramycin resistance cassette in the chromosome of *S. coelicolor* 296

B.3 Null mutants of *sco5718* and *sco2529* do not show developmental or antibiotic defects 298

B.4 Δ*sco2529* and Δ*sco5718* mutants show no change in response to osmotic shock 300

B.5 Confirmation of replacement of the wild type *sco2529/scc117.02* gene with the apramycin resistance cassette in the chromosome of *S. coelicolor* 304

D.1 BldD purified by nickel affinity chromatography 310

List of Tables

Page

Chapter 2

2.1 *E. coli* and *Streptomyces* strains used in this study 36
2.2 Previously developed vectors used in this study 37
2.3 Plasmids generated in this study 39
2.4 Oligonucleotides used in this study 52

Chapter 3

3.1 Potential *bldD* target genes identified by bioinformatic analysis of the *S. coelicolor* genome 88
3.2 BldD-induced changes in gene expression of known and newly identified *bldD* target genes 111
3.3 Description of known and newly identified BldD target genes 112

Chapter 4

4.1 Updated BldD binding sequences for known targets 248
4.2 BldD binding sequences, arranged by strength of binding 251

Appendix

A.1 BldD-induced changes in gene expression of known and putative *bldD* target genes 282
A.2 Description of known and putative BldD target genes 284

List of Abbreviations

α	Alpha or anti (when referring to antibodies, i.e. α-BldD)
Δ	Delta or deletion
γ	Gamma
λ	Lambda bacteriophage
σ	Sigma factor
Ω	Omega or ohm (unit of resistance)
A	Adenine
Å	Angstrom (10^{-10} m)
ABC	ATP-binding cassette
ATP	Adenosine triphosphate
bp	Base pair
C	Cytosine
cAMP	Cyclic adenosine monophosphate
CDA	Calcium dependent antibiotic
cDNA	Complementary DNA
Chaplin	*S. coelicolor* hydrophobic aerial protein
ChIP	Chromatin immunoprecipitation
cpm	Counts per minute
Da	Dalton
ddNTP	Dideoxynucleoside triphosphate
DEPC	Diethyl pyrocarbonate
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTP	Deoxynucleotide
DTT	Dithiothreitol
ECF	Extracytoplasmic function
EDTA	Ethylenediaminetetraacetic acid
EGFP	Enhanced green fluorescent protein
EMSA	Electrophoretic mobility shift assay
G	Guanine
glc	Glucose
GTP	Guanosine triphosphate
HEPES	Hydroxyethylpiperazinethansulphonic acid
IP	Immunoprecipitation
IPTG	Isopropyl β-D-thiogalactopyranoside

kb	kilobase
K_D	Dissociation constant
kDa	kiloDalton
LB	Lysogeny broth
μF	microFarad (unit of capacitance)
MM	Minimal medium
Mmy	Methylenomycin
mRNA	Messenger RNA
MS agar	Mannitol soy flour agar
MW	Molecular weight
NaTCA	Sodium-trichloroacetate
nt	Nucleotide
OD	Optical density
ORF	Open reading frame
PAG	Polyacrylamide gel
PBS	Phosphate buffered saline
PCD	Programmed cell death
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
ppGpp	Guanosine tetraphosphate
R2YE	Sucrose yeast extract medium
RNA	Ribonucleic acid
RNase	Ribonuclease
rpm	Revolutions per minute
rRNA	Ribosomal RNA
Sap	Spore associated protein
SARP	*Streptomyces* antibiotic regulatory protein
SDS	sodium dodecyl sulphate
SELEX	Systematic evolution of ligands by exponential enrichment
SSC	Standard saline citrate
T	Thiamine
T_d	Denaturation temperature
T_m	Melting temperature
TAE	Tris-acetate-EDTA buffer
TBE	Tris-borate-EDTA buffer
TE	Tris-EDTA buffer
tRNA	Transfer RNA

TSB	Trypticase soy broth
tsp	Transcription start point
tsr	Thiostrepton resistance gene
U	Unit (measure of enzyme activity)
UV	Ultraviolet
V	Five or volts
v/v	Volume per volume
w/v	Weight per volume
X-gal	5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside
YEME	Yeast-extract malt-extract medium
YT	Yeast extract tryptone medium

Chapter 1:

Introduction

Chapter 1. Introduction

1.1 General overview

The actinomycetes are a group of Gram positive organisms characterized by their mycelial growth and the high G+C content of their genomes. A key family of this group is the Streptomycetaceae; within this family, members of the genus *Streptomyces* are best known for their ability to produce a wide array of medically and industrially important compounds (secondary metabolites), including antibiotics, antifungal agents (i.e. nystatin), antitumour agents (i.e. daunorubicin), and antiviral, antiparasitic, insecticidal and herbicidal compounds, comprising over two thirds of the naturally occurring biologically active molecules (Challis and Hopwood, 2003; Korn-Wendisch and Kutzner, 1992). The members of this genus are non-fastidious soil-dwellers, and are able to degrade complex and "recalcitrant" substances such as chitin, keratin, and various aromatic compounds. As such, they play an important role in the decomposition of organic matter. They sporulate, and although the spores (also called arthrospores) are not as hardy as endospores, they certainly allow survival through relatively hostile conditions (Korn-Wendisch and Kutzner, 1992).

Streptomyces are obligate aerobes and prefer a neutral to alkaline pH. Like other soil bacteria, they are found more frequently in some sub-regions of the soil than others. Although there are some species capable of causing diseases, most are non-pathogenic. The most important pathogen is *Streptomyces scabies* which causes "common scab" in potatoes and sugar beets. *Streptomyces somaliensis* is the only species pathogenic to humans (Korn-Wendisch and Kutzner, 1992).

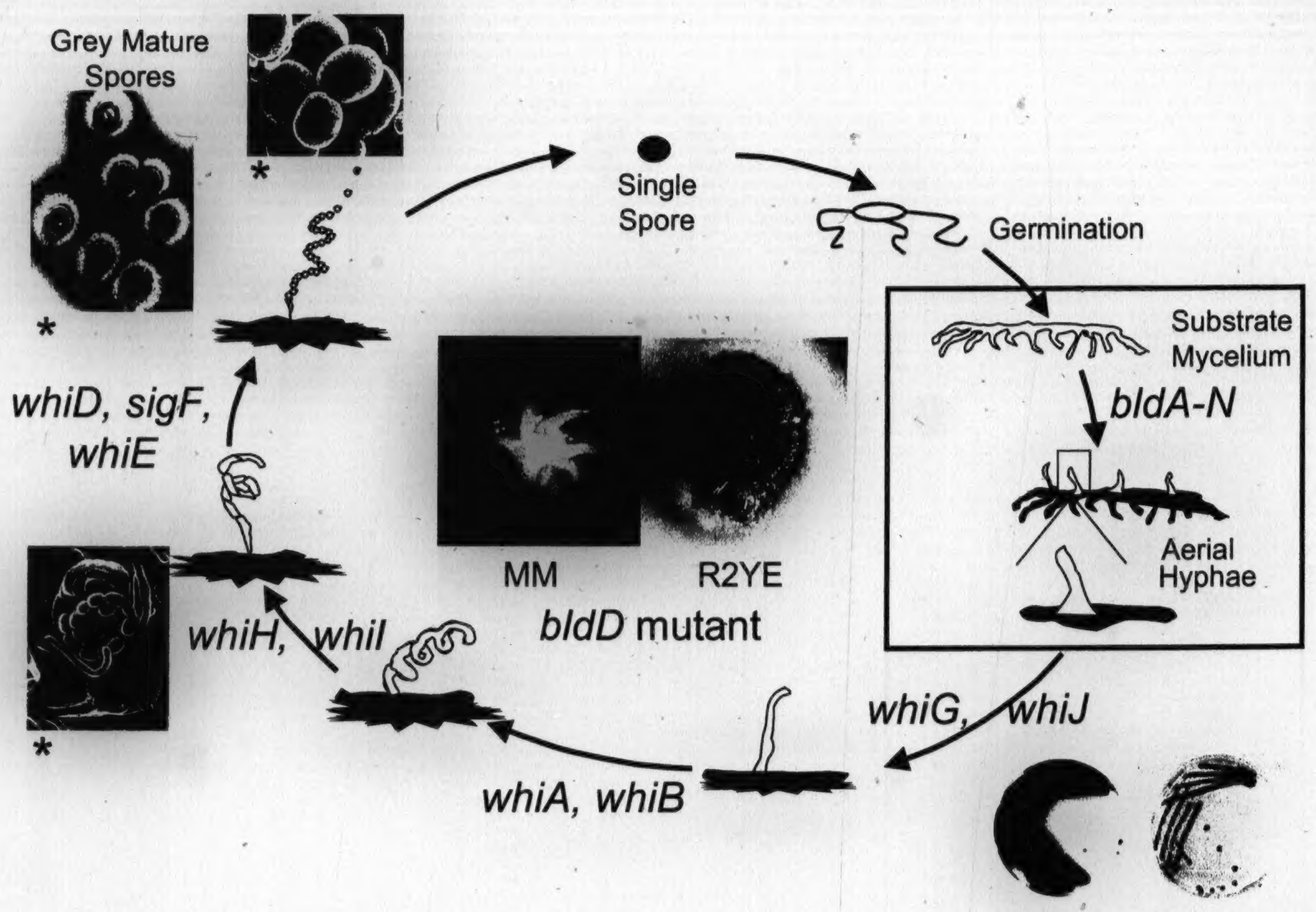
In addition to their production of secondary metabolites, *Streptomyces* spp. are known for their complex life cycle (see Fig. 1.1). Atypically for bacteria, they undergo morphological differentiation to form colonies that superficially resemble fungi. At the 'beginning' of this cycle, a spore germinates to give rise to a mat of substrate mycelium. This vegetative growth consists of long, branched, multinucleoidal cells separated by occasional septa. Following induction by a signal presumed to be nutritional stress, aerial hyphae are erected, breaking the surface tension at the air-water interface of the colony. This is aided by breakdown of accumulated glycogen and trehalose granules (carbon and energy stores), which provides the necessary turgor (Plaskitt and Chater, 1995). The aerial hyphae give the colonies a typical fuzzy or velvety appearance. Erection of the aerial mycelia is accompanied by cannibalization of the substrate mycelium (providing nutritional support for the aerial growth) and production of secondary metabolites (Chater, 1998; Miguelez *et al.*, 1999). Among the secondary metabolites are antibiotics which are thought to be of use in protecting the colony from bacteria and fungi attracted by nutrients released during the breakdown of the substrate mycelium. The aerial hyphae coil and septate to form 50 to 100 unigenomic prespores approximately 1 μm long. In a given hypha, all of the septal divisions form simultaneously. Layers of peptidoglycan are laid down, causing the typical rounding out and thickening of the walls.

Figure 1.1

The *Streptomyces coelicolor* life cycle. Germination of a spore gives rise to a vegetative substrate mycelium. An unknown signal, presumed to be nutritional stress, then stimulates aerial hyphae formation. Aerial growth is mediated through the action of a number of genes, in particular the *bld* genes which are also necessary for the production of antibiotics. Sporulation in the aerial hyphae is regulated by the *whi* genes, which collectively direct coiling and septation of the hyphae and maturation of the spores. Mature spores contain a grey pigment and give the colony its characteristic fuzzy or velvety appearance.

The two pigmented antibiotics produced by *S. coelicolor* are illustrated in the bottom right; actinorhodin (blue) diffuses into the medium surrounding the producing colonies while undecylprodigiosin (red) remains cell-associated.

* Photos courtesy of The Sanger Institute: *S. coelicolor* genome project: http://www.sanger.ac.uk/Projects/S_coelicolor/ and http://www.jic.bbsrc.ac.uk/science/molmicro/Strept.html/ .


Grey Mature
Spores
*
*
Single
Spore
Germination
Substrate
Mycelium
bldA-N
Aerial
Hyphae
whiG, whiJ
whiA, whiB
whiH, whiI
*
whiD, sigF,
whiE
MM
R2YE
bldD mutant

Finally, a grey spore pigment is synthesized using the stored glycogen for precursors and the spores are released (Chater, 1998).

The study of morphological differentiation and secondary metabolite production has been most extensive in *Streptomyces coelicolor* A3(2). This organism is genetically the most characterized and produces at least four antibiotics: actinorhodin and undecylprodigiosin, which are pigmented, and the less-studied calcium-dependent antibiotic and methylenomycin. None of the four have any clinical value, which has allowed much greater collaboration and sharing of information between research groups. The genome of *S. coelicolor* has been sequenced (Bentley *et al.*, 2002), and this sequencing has allowed an even more rapid increase in the body of knowledge regarding this organism.

1.2 The *whi* genes: regulators of sporulation

Much is already known about morphological development and production of the secondary metabolites in *S. coelicolor*. In the late 1970s, one class of genes was shown to be involved strictly in sporulation in *S. coelicolor*. The so-called *whi* (white) genes are named for their characteristic mutant phenotype caused by the absence of the grey spore pigment (Chater, 1972). More recently, some of the *whi* genes were isolated for other characteristics or for homology to other known genes, and were only later found to have a white null phenotype.

Production of the grey spore pigment is the last stage in spore maturation, thus strains with mutations in genes involved in sporulation have a white aerial mycelium instead of the wild type grey. It is important to note that *whi* genes are involved in various stages of sporulation but not in formation of the aerial mycelium (see later for a discussion of the *bld* genes). A large number of *whi* genes have been characterized: *whiA, B, D, E, G, H, I, J, L, M,* and O, as well as *sigF, ftsZ, samR* and the genes encoding the SsgA-like proteins (SALPS). Although very little is known about *samR* (Tan *et al.*, 2002), the rest will be discussed in some detail (see later). The *whi* genes are typically divided into the "early" and "late" genes, determined by the stage at which they act. "Early" genes are needed for the formation of sporulation-specific septa and "late" genes are necessary for maturation of the spores and synthesis of the spore pigment (Kelemen *et al.*, 1996). Converging cascades have been proposed for the *whi* genes, with *whiA/B* and *whiG* exerting effects on *sigF*, but not on each other. Although explained in more detail below, the basis for this proposal is that *sigF* transcription is abolished in both *whiB* and *whiG* mutants, but transcription of *whiB* is independent of *whiG* and vice versa. Thus, two converging pathways, one including *whiG* and the other involving *whiB*, explain these data better than a single, linear one (Chater, 1998).

whiG is one of the earliest acting *whi* genes. The phenotype of a *whiG* disruptant is quite severe, with long, straight aerial hyphae that have vegetative mycelial spacing in the septa (multiple genomes within each compartment) and a complete failure to assemble FtsZ rings (see below); this indicates that *whiG* is necessary for the commitment of the hyphae to sporulation, rather than vegetative growth (Flardh *et al.*, 1999). *whiG* encodes a sigma factor which closely resembles σ^D of *Bacillus subtilis* and σ^{FliA} of *Salmonella typhimurium*, both of which are involved in motility and chemotaxis, not sporulation (Tan and

Chater, 1993; Tan *et al.*, 1998). Surprisingly, *whiG* transcription is not temporally regulated, and transcripts are present at all time points during colony development (Elliot *et al.*, 2001; Kelemen *et al.*, 1996). Given that σ^{WhiG} directs the commitment to sporulation instead of vegetative growth, one might expect it to be expressed just before the initiation of sporulation. This points to post-translational regulation of its activity, as is seen in *S. typhimurium*.

σ^{FliA} of *Salmonella typhimurium* is an alternative sigma factor that controls transcription of the late flagellar genes, including the genes for motility (*mot*) and chemotaxis (*che*). These genes (the σ^{FliA} targets) have a particular conserved sequence in their promoter regions instead of the canonical promoter sequence (Ohnishi *et al.*, 1990). Although itself expressed along with the early flagellar genes, σ^{FliA} is prevented from prematurely activating its targets by the action of FlgM, an anti-sigma factor. FlgM acts by binding to σ^{FliA} and either preventing it from interacting with RNA polymerase in the first place, or by actively dissociating any formed RNA polymerase/σ^{FliA} complexes. Repression is relieved when FlgM is pumped out of the cell and the concentration becomes too low to effectively bind σ^{FliA} (Karlinsey *et al.*, 2000). When *whiG* is introduced into *Salmonella* it is regulated by FlgM (Kelemen *et al.*, 1996), suggesting that there is a homologue of *flgM* in *S. coelicolor*, however examination of the genome sequence has found no obvious candidates (Chater, 2001). The phenotype resulting from *whiG* overexpression is one of hypersporulation in the aerial hyphae and ectopic sporulation in the vegetative mycelia, suggesting that the ratio of WhiG to a putative regulator such as a FlgM-like anti-sigma factor could be an important feature of controlling the timing of sporulation (Kelemen *et al.*, 1996). Several targets of σ^{WhiG} are known, including *proX* (a gene involved in glycine betaine transport) (Tan *et al.*, 1998), *whiI* (Ainsa *et al.*, 1999) and *whiH* (Flardh *et al.*, 1999).

whiA, another early *whi* gene, has homologues in all sequenced Gram positive bacteria, regardless of G+C content (Ainsa *et al.*, 2000). Still, its function is unknown. The *whiA* mutant phenotype is characterized by abnormally long, tightly coiled aerial hyphae with virtually no septa. Although this phenotype clearly places *whiA* after *whiG* in the *whi* gene cascade (Fig. 1), the transcription of *whiA* does not depend on σ^{WhiG}. The wild-type transcript is temporally regulated, with transcription levels increasing sharply at the beginning of sporulation. Self-regulation has been proposed, as temporal regulation of *whiA* transcription is abolished in *whiA* disruption mutants that fail to express WhiA protein (Ainsa *et al.*, 2000).

The mutant phenotype of *whiB* is indistinguishable from that of *whiA*. Again like *whiA*, the function of *whiB* is unknown (Molle *et al.*, 2000), although WhiB has been proposed to be a transcriptional activator because of the presence of a C-terminal helix-loop-helix motif (Davis and Chater, 1992). *whiB*-like genes (*wbl* genes) have been found in all free-living actinobacteria, but not in any other organisms. The products of these genes belong to a family of proteins which feature four conserved cysteine residues and a short tryptophan- and glycine-rich section (Chater and Chandra, 2006). In the mycobacteria, these are often called the *whm* (*w**hiB* homologues of mycobacteria) genes (Soliveri *et al.*, 2000). WhiB3, a WhiB homologue in *Mycobacterium smegmatis*, has been shown to interact directly with the principle σ factor of this organism, giving a possible clue as to the function of the rest of the WhiB homologues (Chater and Chandra, 2006).

whiD is a member of the same family of genes as *whiB*, a family so far found only in actinomycetes. One feature of this family is the absolute conservation of four specific cysteine residues proposed to act as a ligand for a metal cofactor (Molle *et al.*, 2000). WhiD has been shown to bind a [4Fe-4S] cluster *in vitro* (Jakimowicz *et al.*, 2005). The phenotype of a *whiD* null mutant is a severe malformation of the spores. There are many extra septa, resulting in irregular sized spores and minicompartments with no DNA. The spores are also heat sensitive and lyse extensively due to the thin walls. The thinness of the walls stands in contrast to the dramatic thickening between the spores. *whiD* is transcribed from two temporally regulated promoters with similar timing to the late *whi* genes *sigF* and *whiE*, and is not transcribed in submerged culture where *S. coelicolor* does not sporulate (Molle *et al.*, 2000).

WhiH is a member of a large family of bacterial regulatory proteins, and shows most similarity to repressors such as GntR of *Bacillus subtilis*. The members of this family are responsive to carboxylate-containing intermediates in carbon metabolism, such as gluconate (Fisher and Sonenshein, 1991). The transcription of *whiH* is developmentally regulated; that is, it increases when the aerial mycelium is formed and peaks when the spores are first detectable. It is potentially self-regulating, as transcription increases in *whiH* disruption mutants

(Ryding *et al.*, 1998). WhiH may also regulate *whiE* and *sigF* as transcription of these genes is significantly reduced in a *whiH* mutant (Ainsa *et al.*, 1999). *whiH* mutants are pale grey rather than white, possibly due to the reduced transcription of *whiE*, and produce few sporulation septa (Ryding *et al.*, 1999). As stated above, *whiH* is a direct target of σ^{WhiG} (Ryding *et al.*, 1998).

whiI, also a direct target of σ^{WhiG}, has a similar transcription pattern and mutant phenotype to *whiH*, although mutant colonies are white, not pale grey. *whiI* is thought to encode an atypical response regulator (Ainsa *et al.*, 1999; Chater, 2001) that appears to be a member of the FixJ subfamily of two component response regulators, except that it lacks two typical conserved residues found within the phosphorylation pocket of FixJ subfamily proteins - an aspartate and a lysine residue (Ainsa *et al.*, 1999). The conserved aspartate residues within the phosphorylation pocket are responsible for transfer and acceptance of the phosphate group, and through this, for regulation of the activity of the protein. As well, none of the chromosomal genes near *whiI* encode kinases, which is expected of a two-component regulatory gene cluster and is certainly seen with all other known members of the FixJ family of regulators (Hakenbeck and Stock, 1996). To date, no cognate histidine kinase has been found. Together, these suggest that WhiI is not regulated by phosphorylation. As stated above, *whiI* mutants share phenotypic characteristics with *whiH* mutants. Specifically, a *whiI* null mutant has loosely coiled hyphae, no spores and continuous (uncondensed) DNA along the length of the hypha. As with *whiH*, transcription of *whiI* is temporally controlled, and regulated by σ^{WhiG}.

Although there is no evidence for direct interaction between WhiH and the *whiI* promoter or the reverse, *whiH* and *whiI* appear to negatively cross-regulate; transcription of *whiH* is increased in a *whiI* mutant, and vice versa (Ainsa *et al.*, 1999).

Aside from σ^{WhiG}, the only other sigma factor known to act specifically in *S. coelicolor* sporulation is the sigma factor encoded by *sigF*. σ^{F} acts much later than σ^{WhiG} in the sporulation cascade (see Fig. 1.1) and although it is involved in morphological differentiation, it shows the most similarity to σ^{B} of *Bacillus subtilis,* which is involved in general stress response. Spores in a *sigF* mutant are smaller and deformed, more sensitive to detergents and thinner walled. The DNA within the spores is diffuse and uncondensed (Potuckova *et al.*, 1995). The transcription of *sigF* depends indirectly on σ^{WhiG}; σ^{WhiG} does not direct transcription from the *sigF* promoter. *sigF* transcription also depends indirectly on *whiA, -B, -H, -I* and *-J* (Kelemen *et al.*, 1996). σ^{F} itself directs the transcription of one promoter in the *whiE* locus (Kelemen and Buttner, 1998).

whiE, required for production of the grey spore pigment, is the latest acting of the *whi* genes. It is a complex locus of at least 8 open reading frames (ORFs) that closely resemble the components of type II polyketide synthases. Transcription is first seen when sporulation is just detected, then disappears by the end of sporulation. This transcription is completely dependent on *whiA, B, G* and *I*, and partially dependent on *whiH* and *J*. There are two divergent promoters, *whiEp1* for ORFs I-VII, and *whiEp2* for ORF VIII. *whiEp2* is dependent on σ^F, but it is not known which regulatory proteins control transcription from *whiEp1* (Kelemen and Buttner, 1998).

SsgA was the first of a unique family of regulatory proteins to be characterized - the SsgA-like proteins (SALPs) (van Wezel and Vijgenboom, 2004). The SALPs together affect each of the major steps in sporulation, affecting the synthesis or breakdown of peptidoglycan (Noens *et al.*, 2005). Each of these proteins has an important role in sporulation, ranging from sporulation-specific cell division (SsgA and SsgB) to spore wall synthesis (SsgD). The genes for all seven SALPs (*ssgA*-*G*) have been disrupted to determine their role in sporulation (Keijser *et al.*, 2003; Noens *et al.*, 2005; van Wezel *et al.*, 2000), but only *ssgA* has been more thoroughly studied. In *S. coelicolor*, *ssgA* is regulated primarily by SsgR, an IclR-type regulator. SsgR activates *ssgA* transcription in a manner independent of *whiA, whiB, whiG, whiH, whiI* or *whiJ* (Traag *et al.*, 2004). This is very different than the regulation of the orthologue in *Streptomyces griseus*, where *ssgA* is activated by AdpA (key activator of genes involved in development and secondary metabolism), and deletion of the *ssgR* orthologue has no effect on *ssgA* transcription in *S. griseus* (Yamazaki *et al.*, 2003). Disruptants of *ssgA* in *S. coelicolor* produce aerial hyphae on solid media but do not sporulate, and they overproduce actinorhodin (van Wezel *et al.*, 2000). Overexpression of *ssgA* causes increased septation and reduced branching in liquid culture, resulting in spore-like compartments under conditions in which *S. coelicolor* does not sporulate (van Wezel *et al.*, 2000). This increased septation also causes increased fragmentation of the mycelia and reduced clumping, a feature useful for industrial production of secondary metabolites (van Wezel *et al.*, 2006).

ftsZ is essential for proper septum formation during sporulation in *S. coelicolor*. Although a null mutant of *ftsZ* is viable (unique among bacteria), *ftsZ* mutants do not grow as well as the wild type strain and do not sporulate (McCormick *et al.*, 1994). FtsZ forms ladder-like arrays (Z rings) in aerial hyphae, which define the location of the sporulation septa. Ladder formation is synchronous within a given hypha and extends for distances greater than 100 μm. These formations disappear upon completion of septation (Schwedock *et al.*, 1997). Increased levels of *ftsZ* transcription in sporulating aerial hyphae are a result of upregulation of *ftsZ2p*; this increase in transcription is dependent on *whiA, B, G, H, I* and *J*, although transcription from this promoter is not directed by σ^{WhiG} (Flardh *et al.*, 2000).

1.3 Regulation of antibiotic production

In contrast to the genes that for the most part affect only sporulation (the later stage of morphological differentiation) are the genes that affect only antibiotic production. There are more than twenty pathways for secondary metabolites in *S. coelicolor*, most of them cryptic; many have been revealed by sequencing of the *S. coelicolor* genome (Bentley *et al.*, 2002). Only a subset will be discussed here.

Four antibiotics produced by *Streptomyces coelicolor* A3(2) have been characterized – actinorhodin (Act), undecylprodigiosin (Red), calcium-dependent antibiotic (CDA) and methylenomycin (Mmy). Two of these – Act and Red – are pigmented and pH sensitive. Both are red at low pH, but while actinorhodin is blue at high pH, undecylprodigiosin is yellow. Three of the antibiotics diffuse into the media under normal growth conditions, but Red remains cell-associated, as does actinorhodin at low pH (Bibb, 1996). Production of undecylprodigiosin is seen in one particular region of the substrate mycelium (at the center of the colony, near the surface), indicating another level of morphological differentiation (Chater, 1998). The proposed purpose of antibiotic production is to reduce competition at a time when nutrients are scarce; they may even serve to provide another nutrient source for *Streptomyces* by killing the bacteria and fungi attracted by the sugars, etc. released during lysis of the dying substrate mycelium (Challis and Hopwood, 2003). The genes for all four antibiotics are clustered and at least three are under the control of pathway-specific regulators, besides other regulators isolated for their role in controlling the production of multiple antibiotics. The pathway-specific regulators are transcriptional regulators (usually activators) encoded within an antibiotic gene cluster and which control the expression of transcripts encoded in that same cluster. The known or proposed regulators for the antibiotics of *S. coelicolor* are ActII-ORF4, RedD, and CdaR for actinorhodin, undecylprodigiosin and calcium-dependent antibiotic production, respectively (Bibb, 1996). While it was always believed that the pathway-specific genes only regulated the genes in the cluster in which they are encoded, more recently these genes have been shown to each regulate more than just one antibiotic, and even to affect transcription of higher level regulators, indicating that a revision of the current model will be necessary (Huang *et al.*, 2005).

The best characterized antibiotic in *S. coelicolor* is the blue-pigmented polyketide actinorhodin (Act). Transcription of the *act* cluster is controlled by the pathway-specific positive regulator ActII-ORF4 (Arias *et al.*, 1999), a member of the family of DNA-binding regulatory proteins known as SARPs (*Streptomyces* antibiotic regulatory proteins) (Wietzorrek and Bibb, 1997). Other SARPs include the pathway-specific regulators of both undecylprodigiosin biosynthesis (RedD) and calcium-dependent antibiotic (CdaR). As well, AfsR, a pleiotrophic regulator of both actinorhodin and undecylprodigiosin production in *S. coelicolor*, is a SARP (Wietzorrek and Bibb, 1997).

The other pigmented antibiotic in *S. coelicolor* is the tripyrrole antibiotic undecylprodigiosin (Red), which is the major component of a mixture of similar compounds (prodigionines) made by *S. coelicolor* (Tsao *et al.*, 1985). Red production may be a sink for excess proline in the cell; this amino acid appears to be a precursor for undecylprodigiosin production (Hopwood *et al.*, 1994). As stated above, the SARP controlling Red production is RedD, however regulation of Red production is more complex than that of actinorhodin production. In addition to the cluster-encoded SARP, Red production is also dependent on another *bldA*-dependent (see section 1.4) pathway-encoded regulator, RedZ. RedZ is necessary for the transcription of RedD (White and Bibb, 1997). RedZ is not itself a SARP but resembles the response regulators of two-component response regulator systems, though without the conserved residues in the phosphorylation pocket (Guthrie *et al.*, 1998).

Not much is known about CDA, the calcium-dependent antibiotic. As the name implies, it is produced only in the presence of high concentrations of Ca^{2+}. CDA is a lipopeptide with a non-ribosomally derived peptide moiety and a six-carbon hydroxylated fatty acid at its N-terminus (Chong *et al.*, 1998). CDA is chromosomally encoded and *absA*, one of the pleiotropic regulators of *S. coelicolor* (discussed later), is encoded between the structural (peptide synthetase) genes and *cdaR*, the putative SARP-like regulator of the *cda* cluster (Anderson *et al.*, 2001; Ryding *et al.*, 2002).

Methylenomycin A (Mmy) is encoded on the giant linear plasmid SCP1 and is the only antibiotic among the actinomycetes known to be plasmid-encoded (Bibb, 1996). Mmy is the first antibiotic to be produced during the growth cycle of *S. coelicolor*, appearing before entry into stationary phase during a period of high glucose consumption and medium acidification (Hobbs *et al.*, 1992). Undecylprodigiosin is produced next, at the beginning of stationary phase, then actinorhodin six to eight hours later (Strauch *et al.*, 1991).

A number of genes have been studied because their deletion causes defects in production of multiple antibiotics; they were often isolated because their overexpression stimulates actinorhodin production in *Streptomyces lividans*, a strain which does not normally produce this antibiotic. These include the *absA* locus, encoding a two component regulatory pair (Brian *et al.*, 1996), *absB*, encoding an RNase III homologue (Price *et al.*, 1999), *afsR*, encoding a regulatory protein with N-terminal similarity to the SARPs (Horinouchi, 2003), and the *afsQ* locus, encoding a two component regulatory pair with an associated putative accessory protein (Ishizuka *et al.*, 1992), (M. Hutchings, posted on ScoDB - http://streptomyces.org.uk/). As these genes are not central to the thesis, they will not be discussed in detail.

1.4 The *bld* genes: regulators of morphological differentiation and antibiotic production

There is a hierarchy of regulatory genes controlling antibiotic production in *S. coelicolor*. As already mentioned, antibiotic production is controlled at the level of pathway-specific regulators, which are in turn controlled by more global regulators controlling more than one antibiotic pathway. Lastly, there is yet another level of regulation that involves genes encoding products that control both secondary metabolite production and morphological differentiation. This class of regulators includes the *bld* genes, which are named for the shiny or bald appearance of mutant colonies (in contrast to the fuzzy or velvety appearance of wild-type colonies expressing aerial hyphae). In addition to the "bald" phenotype, which indicates involvement in the erection of aerial hyphae and for which the mutants were named, *bld* mutants are also pleiotropically blocked in antibiotic production. For example, *bldA, -B, -D, -G,* and *-I* are involved in regulation of all four antibiotics (Champness, 1988; Merrick, 1976). The control of antibiotic production makes understanding the *bld* genes very important; increased knowledge of the mechanism by which the *bld* genes act has the potential to allow more efficient production of antibiotics and other industrially important secondary metabolites.

To date, 13 different *bld* genes are known that are named as such, besides others described for the phenotype of partial disruptants, or for homology to known genes in other systems. At the moment, *bldA, B, C D, G, H, K, M* and *N* have been cloned (Chater and Chandra, 2006), and as more genes involved in development are cloned and characterized, the mechanisms by which aerial hyphae formation and antibiotic production are regulated will become clearer. *bld* mutants are either unable to make aerial mycelium, or have mycelia that lie prostrate on the colony surface (*bldA* and *D*) (Champness, 1988; Elliot *et al.*, 1998). These prostrate mycelia are presumed to be aerial mycelia without the turgor or surface hydrophobicity necessary to break the surface tension and extend into the air. The inability of some *bld* mutants to form or erect aerial hyphae is complemented by growth on poorly used carbon sources, such as mannitol (Champness, 1988; Merrick, 1976). Exceptions to this rule include mutants of *bldB*, *bldM* and *bldN* (Bibb *et al.*, 2000; Molle and Buttner, 2000; Pope *et al.*, 1996). In contrast to the morphological phenotype, complete complementation of the antibiotic-negative phenotype by growth on minimal medium with mannitol as the sole carbon source is only seen in *bldH* mutants (Champness, 1988).

One of the most interesting features of the *bld* mutants is that they exhibit extracellular complementation; that is, when a pair of mutants are grown next to each other, one will develop a fringe of aerial mycelia on the side of the growth patch nearest the other. This complementation is always unidirectional, with one mutant acting as the donor and the other as the recipient. A cascade has been developed from the specific pattern of the complementation: *bldJ* < *bldK, bldL* < *bldA, bldH* < *bldG* < *bldC* < *bldD, bldM* (Chater, 2006; Molle and Buttner, 2000; Nodwell *et al.*, 1996; Willey *et al.*, 1993). Each *bld* mutant can complement all of the *bld* mutants to its left, but none of those to its right. Genes separated by a comma are considered to be in the same complementation group. Based on this observed extracellular complementation, it has been proposed that aerial hyphae formation is initiated by a signalling cascade involving at least five extracellular signals, symbolized by <. From this theory, one would expect the *bld* genes to encode extracellular signals or their receptors. Two of the *bld* genes (*bldJ* and *bldK*) fit this expectation, but the rest must have, at best, indirect roles in the extracellular signalling (Chater, 1998).

The first *bld* gene to be cloned and characterized was *bldA*. It encodes the only tRNA capable of translating the rare UUA codon in the high G+C, *S. coelicolor* mRNA. TTA codons are found in several genes upregulated during morphogenesis including the pathway-specific regulators of actinorhodin (*act*II-ORF4) (Fernandez-Moreno *et al.*, 1991) and undecylprodigiosin (*redZ*) biosynthesis (Guthrie *et al.*, 1998), and 145 other open reading frames identified by the *Streptomyces coelicolor* genome project (Bentley *et al.*, 2002; Kim *et al.*, 2005). The transcription of *bldA* depends on *bldI* (Leskiw and Mah, 1995) and is temporally regulated, with the mature tRNA accumulating markedly later in growth (Leskiw *et al.*, 1993). Upregulation of TTA-containing transcripts is one level of regulation, but clearly translation of UUA codons is also temporally regulated through *bldA*.

bldB is involved in regulation at a transcriptional, rather than at a translational, level. BldB is a small (98 amino acid), highly charged protein predicted to be a transcriptional regulator (Pope *et al.*, 1998). *bldB* has been shown to regulate its own transcription (Pope *et al.*, 1998). BldB has a helix-turn-helix motif near its C-terminus and a tyrosine residue important for its function, as determined by mutation studies (Pope *et al.*, 1998). BldB dimerizes, but dimerization is separate from function (Eccleston *et al.*, 2002; Eccleston *et al.*, 2006). Overexpression of *bldB* in wild type *S. coelicolor* causes a *whi* phenotype; aerial hyphae are deficient in sporulation septation. In contrast, overexpression of a functionally defective but dimerization-competent form of BldB caused precocious sporulation in the wild type strain, though not in *bldK, ramR, whiA, whiB, whiG, whiH, sigF and ssgB* mutants. Together, these results suggest that BldB is a negative regulator of development (Eccleston *et al.*, 2006). Unlike most *bld* mutants, *bldB* mutants are not phenotypically rescued by growth on the poor carbon source mannitol. In *bldB* mutants, deregulation appears to be global and several promoters involved in carbon metabolism are affected, suggesting that *bldB* regulates carbon source use rather than directly regulating developmental genes (Pope *et al.*, 1996). There are many homologues of *bldB* in *S. coelicolor*, the more recently sequenced *S. avermitilis* genome, and in a number in other filamentous actinomycetes, but none in the non-filamentous actinomycetes or any other group of bacteria (Eccleston *et al.*, 2006).

bldC was identified in the original screen for *bld* mutants, and has recently been shown to encode a DNA binding protein of a novel family of small DNA-binding proteins found in both Gram positive and Gram negative bacteria (Hunt *et al.*, 2005; Merrick, 1976). BldC has a MerR-type DNA binding domain but does not have the C-terminal effector binding or dimerization domains typical of the MerR family (Hunt *et al.*, 2005). Depending on the genetic background, *bldC* mutants are either bald on all media, or bald on minimal media and severely delayed in development on rich media. Although *bldC* mutants do not produce actinorhodin due to a transcriptional dependence of *actII-orf4* on *bldC*, they do produce undecylprodigiosin. The latter is produced despite a partial dependence of *redD* on *bldC*; *bldC* is necessary for maintenance of *redD* transcription during aerial hyphae formation and sporulation (Hunt *et al.*, 2005). Transcription of *bldC* does not depend on any of the known *bld* genes (Hunt *et al.*, 2005).

bldG was originally described by Dr. Wendy Champness in 1988. Defects in morphological differentiation were found to be carbon-source dependent; growth on minimal medium with mannitol as the sole carbon source rescued the aerial mycelium formation (Champness, 1988). *bldG* encodes a putative anti-anti-sigma factor. It is transcribed both as a mono- and as a polycistronic mRNA, the latter with a downstream open reading frame (*orf3*) which shows homology to anti-sigma factors. Anti-sigma factors interact with sigma factors to prevent their interaction with RNA polymerase core enzyme; anti-anti-sigma factors disrupt the interaction and bind to the anti-sigma factors, preventing anti-sigma/sigma factor complexes. Surprisingly, there is no sigma factor encoded at the same locus;

this could indicate regulation of multiple sigma factors. There are three promoters upstream of *bldG*, two of which are induced during the formation of aerial mycelium and antibiotics (Bignell *et al.*, 2000; Bignell *et al.*, 2005). *bldG*, although a regulator of transcription, does not regulate its own transcription (Bignell *et al.*, 2000). The BldG protein is regulated by phosphorylation, as is the case for proteins known to function as anti-sigma factors (Bignell *et al.*, 2003). The transcription of *bldN* is dependent on *bldG* (Bibb *et al.*, 2000).

bldH is a gene through which *bldA* exerts some of its effects on morphological differentiation (Nguyen *et al.*, 2003; Takano *et al.*, 2003). The *adpA* gene in the *S. coelicolor bldH109* mutant (Champness, 1988) has a frameshift mutation suggesting that *adpA* is the *bldH* gene (Nguyen *et al.*, 2003; Takano *et al.*, 2003). Transcription of *adpA* is developmentally regulated, peaking just before aerial hyphae formation and decreasing several hours thereafter (Nguyen *et al.*, 2003). The wild type *S. coelicolor adpA* gene has a TTA codon, indicating *bldA* regulation of *adpA*. Mutation of the *adpA* TTA codon to CTC or TTG allows partial complementation of the deficiency in aerial hyphae formation of a *bldA* mutant, indicating that *bldH* is responsible for the *bldA* mutant phenotype (Nguyen *et al.*, 2003; Takano *et al.*, 2003). Regulation of the *S. coelicolor adpA* gene is quite different than that of its close homologue (84.4% identity) in *S. griseus*; *adpA* transcription in *S. coelicolor* does not depend on γ-butyrolactones such as A factor, which is critical to *S. griseus* development (Takano *et al.*, 2003).

In contrast to several of the genes discussed above, *bldJ* (formerly *bld261*) is responsible for the production of an extracellular signal, most likely a peptide, that restores aerial hyphae formation to *bldJ* mutants in nanomolar concentrations. *bldJ* has not been cloned and attempts to sequence the putative signalling peptide were unsuccessful. An intact *bldK* locus is necessary for import of this signalling molecule; purification of the *bldJ*-dependent signal was possible because it accumulates in the medium in the absence of the BldK transporter (Nodwell and Losick, 1998).

bldK is a complex locus encoding an ATP-binding cassette family oligopeptide permease. This transporter is thought to function as an importer of the *bldJ* – dependent extracellular signal (Nodwell *et al.*, 1996) *bldJ* and *bldK* are the only *bld* genes discovered to date that fit directly with the model of a *bld* gene-dependent extracellular signalling cascade described earlier.

bldN was originally isolated as *whiN* in a screen for *whi* mutants, but a null mutant displayed a *bld* phenotype. The phenotype of *bldN* null mutants indicates that *bldN* is involved only in morphological differentiation (not antibiotic production). In this respect, *bldN* represents a potential branch point in a signalling cascade. *bldN* encodes an extracytoplasmic function sigma factor – that is, it is responsive to environmental signals and affects functions outside of the cytoplasm (Bibb *et al.*, 2000). There is an unusual N-terminal extension in the translated sequence which is cleaved by an as-yet unknown protease (Bibb and Buttner, 2003). This type of extension has been seen in the pro (inactive) forms of *Bacillus subtilis* σ^E and σ^K, but never previously in a *Streptomyces* sigma factor. In *B. subtilis*, the extension is cleaved to activate the sigma factor, allowing it to move freely in the cytoplasm and interact with RNA polymerase and the DNA (Bibb *et al.*, 2000). In *S. coelicolor*, the processing takes place in a development-dependent manner; the processed form of σ^{BldN} is not seen in significant levels until just before the initiation of aerial hyphae formation (Bibb and Buttner, 2003). Transcription of *bldN* occurs from a single promoter and is indirectly dependent on *bldG* and *bldH* (Bibb *et al.*, 2000).

The other *bld* gene in *S. coelicolor* involved solely in morphological differentiation is *bldM* (formerly *whiK*), discovered in the same screen as *bldN*. BldM encodes a response regulator of the FixJ subfamily of two component response regulators, although it is not regulated by phosphorylation (Molle and Buttner, 2000). *bldM* is a direct biochemical target of σ^{BldN} and its transcription depends on *bldN* (Bibb *et al.*, 2000). This is the first example of direct regulation of one *bld* gene by another. Interestingly, while *bldM* fits in the extracellular signalling cascade, *bldN* does not (Chater, 2001).

brgA is a *bld* gene, but was isolated by screening for mutants resistant to 3-aminobenzamide (an inhibitor of ADP-ribosyltransferase). Null mutants show defects in aerial hyphae formation, ADP ribosylation, and pigmented antibiotic production (production of calcium-dependent antibiotic and methylenomycin was not tested). ADP ribosylation is an enzymatic reaction well understood in eukaryotes but little understood in prokaryotes. ADP-ribosyltransferase binds the ADP-ribose moiety of NAD to specific acceptor proteins; this is a form of metabolic regulation. Like *bldB* mutants, *brgA* mutants are bald regardless of carbon source. Interestingly enough, *bldB* mutants are also deficient in ADP-ribosylation, as are *bldC* and *bldH* mutants. The phenotype of *bldH* with respect to this process is the most severe; mutants are completely unable to ADP-ribosylate any proteins at any point during the cell cycle (Shima *et al.*, 1996).

A number of screens have identified new loci important for development. One screen identified 50 new *bld* mutant strains, including a novel *bld* mutant, *bldL*. *bldL* mutants behave like *bldK* mutants but map to a chromosomally distinct location (Nodwell *et al.*, 1999). More screens have been carried out using transposon mutagenesis methods developed for *S. coelicolor* (Bishop *et al.*, 2004; Gehring *et al.*, 2000). Among the genes identified in these screens is *osaB*, necessary for proper development and antibiotic production (Bishop *et al.*, 2004). *osaB* encodes a response regulator predicted to interact with another protein instead of binding DNA. Null mutants are unable to develop properly in the presence of osmotic stress, having no aerial hyphae and overproducing both pigmented antibiotics (Bishop *et al.*, 2004).

As BldD is central to this thesis, a detailed review is in order. BldD is a small (167 residue), DNA-binding protein of the XRE (Xenobiotic Response Element) family (Elliot and Leskiw, 1999; Elliot *et al.*, 2001). The original *bldD* mutant was isolated in the mid 1970s (Merrick, 1976). The mutant strain was bald (had no aerial hyphae) and deficient in production of all antibiotics tested for, though there were malformed hyphae lying prostrate on the colony surface, possibly aerial hyphae without sufficient turgour or surface hydrophobicity to rise into the air (Elliot *et al.*, 1998; Merrick, 1976). Growth on minimal medium with mannitol as the sole carbon source rescues the morphological defect of *bldD* mutants but not the production of antibiotics (Pope *et al.*, 1996) The original strain contained a point mutation changing the amino acid at position 62 from a tyrosine to a cysteine (Elliot *et al.*, 1998). This mutation, which resulted in unstable BldD protein, was originally proposed to interfere with BldD activity through replacement of a phosphorylated residue (Elliot *et al.*, 1998). However, the substitution was later shown to instead disrupt the hydrophobic core of the N-terminal DNA-binding domain of BldD, disrupting the protein structure and resulting in the observed instability of mutant BldD (Elliot, 2000; Kim *et al.*, 2006).

BldD was first shown to bind to its own promoter (Elliot and Leskiw, 1999), then later to the promoters of other target genes. All of the confirmed BldD targets to date are regulators or putative regulators. Two of the known targets are regulators of development; *bldN* encodes an extracytoplasmic sigma factor required for aerial hyphae formation and *whiG* encodes a sigma factor critical for initiation of sporulation (Elliot *et al.*, 2001). A third target, *sigH*, encodes a sigma factor responsive to a number of stresses, though the stress response does not appear to be mediated by BldD (Kelemen *et al.*, 2001). BldD regulates the expression of *sigH* in space as well as time; restriction of *sigH* expression to the aerial hyphae is *bldD*-dependent (Kelemen *et al.*, 2001). The last published target of BldD is *bdtA*, a putative DNA-binding regulator of unknown function (Elliot *et al.*, 2001). A *bldD*-dependent activation of *whiG* and *bdtA* is seen in late growth, though the apparent requirement during late development for *bldD* could be explained by a lack of aerial hyphae in the *bldD* mutant rather than by BldD-mediated activation of these genes in the wild type (Elliot *et al.*, 2001).

A null mutant of *bldD* has been obtained through the complete deletion of the open reading frame derived through the PCR-targeted mutagenesis protocol of Gust *et al.*, 2003 (Elliot *et al.*, 2003b). Visually, the null mutant is identical to the original point mutant (Elliot *et al.*, 2003b). In the same work, BldD was shown to exist primarily as a dimer in solution, with some protein present in monomer or tetramer form, though BldD binds DNA solely as a dimer (Elliot *et al.*, 2003b).

The *bldD* homologue in *S. griseus* has been cloned and the product shows a remarkable degree of sequence identity to BldD in *S. coelicolor*; 164 of 167 amino acids in BldD are identical, and the three surrounding genes sequenced with *bldD* encode proteins that show greater than 90% similarity to their counterparts in *S. coelicolor* (Ueda *et al.*, 2005). A homologue of *bldD* is also present in *S. avermitilis*, with 93% DNA sequence and 98% amino acid sequence identity (Ikeda *et al.*, 2003) (NCBI BLAST, http://www.ncbi.nlm.nih.gov/sutils/genom_tree.cgi) The crystal structure of the *S. coelicolor* BldD N-terminus (BldDN) has been determined to a resolution of 1.8 Å (Kim *et al.*, 2006). BldDN is globular and compact, containing four alpha-helices, including two (α2 and α3) which form a helix-turn-helix motif similar to the DNA-binding domain of the lambda repressor, another XRE family member (Kim *et al.*, 2006). Mutational studies indicated that α3 and the turn region between the two helices in the helix-turn-helix motif (α2 and α3) are particularly important for sequence recognition (Kim *et al.*, 2006). As the search for more targets of BldD regulation was the central purpose of this work, further details can be found in the Results and Discussion sections.

1.5 The *ram* and *chp* genes: surfactants and surface proteins

A much studied group of genes important for development is the *ram* cluster. *ramA, B* and *R* from *S. coelicolor* were found by their ability to cause accelerated aerial mycelium formation (48 vs. 72-96 hours) in the closely related *S. lividans*, and named the *ram* genes (rapid aerial mycelium) (Ma and Kendall, 1994). Although only *ramA*, *ramB* and *ramR* were found in the initial screen, other work found the rest of the cluster - *ramC* and *ramS* (Keijser *et al.*, 2000). The *ram* genes of *S. coelicolor* are arranged as a cluster; the *ramR* gene is convergent to the *ramCSAB* operon. *ramA* and *ramB* are predicted to encode components of an ABC transporter, likely to export the product of *ramS* from the cell (Ma and Kendall, 1994). The product of the *ramR* gene is an "orphan" response regulator; there is no histidine kinase encoded nearby, although there is a kinase-like domain in the N terminus of the protein (Hudson and Nodwell, 2004; O'Connor *et al.*, 2002). The product of the *ramC* gene was initially predicted to be a receptor kinase, based on similarity to serine/threonine protein kinases in the N-terminus (O'Connor *et al.*, 2002), but later predicted to have an entirely different function, important to the processing of the product of *ramS* (Kodani *et al.*, 2004). See below for further discussion.

Deletion of *ramR* or *ramC* causes a severe delay in the production of aerial hyphae, although neither gene is essential for vegetative growth (Nguyen *et al.*, 2002; O'Connor *et al.*, 2002). Conversely, overexpression of *ramR* in various *bld* mutant strains restored aerial hyphae production, although without correcting the defects in antibiotic production associated with the *bld* mutations (Nguyen *et al.*, 2002). RamR acts at least in part through direct activation of the *ramCSAB* operon (O'Connor and Nodwell, 2005); expression of the operon is also dependent on *cprA* and *bldD*, though not on *bldM* or *bldN*. Direct regulation has been shown by RamR binding upstream of *ramC* (O'Connor *et al.*, 2002).

Recently, the product of *ramS* was shown to be cleaved and modified to form SapB, a secreted peptide (surfactant) which acts to reduce the surface tension at the colony surface and allow the aerial hyphae to rise into the air (Kodani *et al.*, 2004; Tillotson *et al.*, 1998). The exact nature of SapB was a mystery for many years; Edman degradation was blocked after the fifth cycle, preventing full sequencing of the peptide (Willey *et al.*, 1991). Examination of the C terminal sequence of RamC independent of the rest of the protein finally unearthed a clue to the nature of SapB; RamC was shown to be a likely Lan cyclase (an enzyme involved in the maturation of lantibiotic peptides). At the same time, sequencing of SapB was attempted after subjecting it to a treatment known to assist in sequencing of lantibiotic peptides. This experiment showed that the N terminal sequence of SapB matched the RamS sequence starting at position 22 (Kodani *et al.*, 2004). SapB was subsequently shown to be a lantibiotic-like peptide, though with no antibiotic activity, that is produced from the initial RamS peptide by a combination of proteolytic cleavage, dehydration of residues, and formation of lanthionine bridges (Kodani *et al.*, 2004). *ramC* expression is limited to the substrate mycelium and very young aerial hyphae

(O'Connor *et al.*, 2002). This is consistent with the role of the operon in producing a surfactant important for the erection of the aerial hyphae.

SapB is not the only surfactant involved in aerial hyphae formation, and there is at least one route to aerial hyphae formation that exists independent of the *ram* genes; some *bld* gene mutants (unable to form aerial hyphae) are able to differentiate when grown on minimal medium with mannitol as the carbon source, while still unable to produce SapB (Champness, 1988). A newly discovered cluster of genes, *ragABKR* (*ram* activated genes), activates a SapB-independent differentiation and sporulation pathway downstream of *ramR* (San Paolo *et al.*, 2006). Unlike disruption of the *ram* genes, mutation of the *rag* operon does not result in a bald phenotype but rather one like a *whiG* mutant - branched, undifferentiated aerial hyphae that resemble the substrate mycelia. Grey pigmentation of the hyphae explains why this cluster was not discovered in screens for *whi* mutants. Simultaneous overexpression of *ragK* (putative histidine kinase) and *ragR* (*ramR* paralogue) rescues not only aerial hyphae formation but also sporulation in a number of *bld* and *ram* gene mutants, though with no concurrent SapB production. This operon is likely responsible for the production of a diffusible signal, exported by the products of *ragAB* (putative ABC transporter), that stimulates differentiation of aerial structures (San Paolo *et al.*, 2006).

Deletion of *ramR*, resulting in no expression of either the *ram* or the *rag* genes still allows aerial hyphae production, although severely delayed (Nguyen *et al.*, 2002). This indicates the presence of other surfactants produced by *S.*

coelicolor. Some of the recently identified chaplins (for *S.* *coelicolor* hydrophobic aerial proteins) (ChpD-H) can also reduce water tension sufficiently to allow erection of aerial hyphae (Claessen *et al.*, 2003). The main role of the chaplins, however, is not so much in reducing surface tension initially as in coating the surface of the emerging aerial hyphae and providing them with a hydrophobic surface which allows continued aerial growth (Claessen *et al.*, 2004; Elliot *et al.*, 2003a). In the absence of all eight chaplins, the few aerial hyphae produced collapse on the colony surface from absorption of water, due to their hydrophilic surfaces (Claessen *et al.*, 2004). In contrast to SapB, the chaplins do not diffuse through the medium, so that extracellular complementation is not seen between wild type and *chp* mutant strains (Claessen *et al.*, 2003; Elliot *et al.*, 2003a). As well, addition of purified chaplins does not rescue the phenotype of *bld* gene mutants (Claessen *et al.*, 2003).

1.6 Thesis objectives

The purpose of this thesis was to find and characterize additional targets of BldD and to determine their role in the morphological and physiological differentiation of *Streptomyces coelicolor*. BldD was previously shown to bind to its own promoter by gel-shift assays, and the inverted repeat bound by BldD was determined by DNase I mapping (Elliot and Leskiw, 1999). Further footprinting assays allowed determination of a consensus binding sequence - AGTgA (n_m) TCACc, where n indicates any base in a sequence of variable length (m) (Elliot *et al.*, 2001). The consensus between the sequences bound in the different targets was very loose and predicting putative targets on the basis of the consensus alone was highly impractical, as discussed later in this thesis. Given the limitations of the SELEX (systematic evolution of ligands by exponential enrichment) procedure previously used (Elliot *et al*., 2001), new methods for finding BldD targets had to be found. Microarray analysis has proven very successful in global screens of gene expression, and was therefore chosen as a method for BldD target identification. The targets identified in this thesis were further evaluated by both *in vitro* and *in vivo* methods to provide convincing evidence that BldD does, in fact, regulate their transcription. In this thesis, I present evidence that *prsI* and *arsI*, genes encoding an anti-sigma factor and an anti-sigma factor antagonist respectively, are direct targets of BldD regulation. I also present evidence that BldD likely regulates the transcription of *sigQ* which encodes a sigma factor of unknown purpose. Some of the other targets found by array analysis will also be listed. Together, these genes present both new puzzles to be solved and an insight into the complex developmental processes of *S. coelicolor*.

Chapter 2:

Materials and Methods

Chapter 2. Materials and Methods

2.1 Bacterial Strains, Plasmids and Growth Conditions

2.1.1 *Streptomyces coelicolor* and *Escherichia coli* strains

Streptomyces coelicolor and *Escherichia coli* strains used in this study are found in Table 2.1.

2.1.2 Plasmids

Previously constructed plasmids used in this study are listed in Table 2.2, and plasmids generated in this study are listed in Table 2.3.

2.1.3 *E. coli* growth conditions

Cells were grown at 30°C or 37°C, as indicated. Liquid cultures were grown in LB broth (1% w/v tryptone, 0.5% w/v yeast extract, 1% w/v NaCl) or SOB medium (2% w/v tryptone, 0.5% w/v yeast extract, 0.05% w/v NaCl) (Sambrook *et al.*, 1989) on a rotating rack. Solid cultures were plated on LB agar (1.5% w/v agar). To maintain plasmids, antibiotics were added as indicated: ampicillin (Sigma) to 100 μg/mL, apramycin (Provel) to 50 μg/mL, chloramphenicol (Aldrich) to 25 μg/mL, kanamycin (Sigma) to 30 μg/mL and spectinomycin (Sigma) to 100 μg/mL. For long term storage, cultures, inoculated from a single colony, were grown for 16 h, then mixed with sterile 40% glycerol

Table 2.1 *E. coli* and *Streptomyces* strains used in this study

Strain	**Genotype**	**Reference and/or source**
S. coelicolor		
M600	prototroph, SCP1$^-$, SCP2$^-$	(Chater *et al.*, 1982), Gift from M. Buttner, John Innes Centre
M145	prototroph, SCP1$^-$, SCP2$^-$, Pgl$^+$	(Lomovskaya *et al.*, 1980), John Innes Centre
916	*hisA1 mthB2 pheA1 strA1* SCP1^{NF} (SCP2*)	(Merrick, 1976), John Innes Centre
1169	*hisA1 mthB2 pheA1 strA1 bldD53* SCP1^{NF} (SCP2*)	(Merrick, 1976), John Innes Centre
Δ*arsI*	prototroph, SCP1$^-$, SCP2$^-$, Δ*arsI::aac(3)IV* (M600 derivative)	This study
Δ*bldD*1	prototroph, SCP1$^-$, SCP2$^-$, Δ*bldD::aac(3)IV* (M600 derivative)	Elliot *et al.*, 2003
Δ*sco5718*	prototroph, SCP1$^-$, SCP2$^-$, Δ*sco5718::aac(3)IV* (M600 derivative)	This study
Δ*sco2529*	prototroph, SCP1$^-$, SCP2$^-$, Δ*sco2529::aac(3)IV* (M600 derivative)	This study
Δ*sigI*	prototroph, SCP1$^-$, SCP2$^-$, Δ*sigI::aac(3)IV* (M600 derivative)	This study
Δ*sigQ*	prototroph, SCP1$^-$, SCP2$^-$, Δ*sigQ::aac(3)IV* (M600 derivative)	This study
E. coli		
BW25113	Δ *(araD-araB)567*, Δ*lacZ4787(::rrnB-4)*, *lacIp-4000(lacI^Q)*, λ$^-$, *rpoS3699(Am)*, *rph-1*, Δ*(rhaD-rhaB)568*, *jsdR514*	(Gust *et al.*, 2003), Plant Bioscience Limited, U.K.
DH5α	F' *supE44* Δ*ladU149 (Φ80lacZΔM15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1*	(Hanahan, 1983), Gibco BRL
ET12567	F$^-$, *dam-13::Tn9*, *dcm-6*, *hsdM*, *hsdR*, *recF143*, *zjj-202::Tn10*, *galK2*, *galT22*, *ara14*, *lacY1*, *xyl-5*, *leuB6*, *thi-1*, *tonA31*, *rspL136*, *hisG4*, *tsx78*, *mtl-1*, *glnV44*	(MacNeil *et al.*, 1992), Gift from D. MacNeil, Merck Sharp and Dohme Research Laboratories
JM109	*endA1*, *recA1*, *gyrA96*, *thi*, *hsdR17*(r_k-, m_k+), *relA1*, *supE44*, λ-, Δ*(lac-proAB)*, [*F'*, *traD36*, *proA$^+$B$^+$*, *lacI^{q}Z*Δ*M15*]	Yanisch-Perron *et al.*, 1985
One Shot® Top 10	F$^-$ *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 ΔlacX74 *rec*A1 *ara*D139 Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G	Invitrogen

Table 2.2 Previously developed vectors used in this study

Plasmid	Features	Antibiotic resistance	Reference and/or source
***E.coli/ Streptomyces* shuttle plasmids**			
pAU181	1.3 kb SphI/XmnI fragment in pSET152	apramycin	Elliot *et al.*, 1998
pIJ6902	pSET152 derivative with the thiostrepton-inducible promoter *ptipA* from *Streptomyces* and a thiostrepton resistance gene (*tsr*)	apramycin thiostrepton	Gift from M. Buttner, John Innes Centre, (Huang *et al.*, 2005)
pSET152	High copy number in *E. coli*, integrates into the *attB* site of *Streptomyces* through the ϕC31 integration system	apramycin	(Bierman *et al.*, 1992), Northern Regional Research Center, Peoria, Ill.
pSETΩ	pSET152 derivative where *aadA* replaces *aac(3)IV*	spectinomycin	Gift from J. Nodwell, (O'Connor *et al.*, 2002)
pHJL400	high copy number in *E. coli*, par$^-$ in *Streptomyces*	ampicillin, thiostrepton	Larson and Hershberger, 1986
***E. coli* plasmids, phagemids and cosmids**		kanamycin carbenicillin	
2SCK8	Supercos 1 derivative containing 38.3 kb of *S. coelicolor* chromosomal DNA, including the *sigQ* gene	kanamycin carbenicillin	Gift from H. Kieser, John Innes Centre, (Redenbach *et al.*, 1996)
SC3C3	Supercos 1 derivative containing 39.2 kb of *S. coelicolor* chromosomal DNA, including the *sco5718* gene	kanamycin carbenicillin	Gift from H. Kieser, John Innes Centre, (Redenbach *et al.*, 1996)
SC9C5	Supercos 1 derivative containing 43.7 kb of *S. coelicolor* chromosomal DNA, including the *bldD* gene	kanamycin carbenicillin	Gift from H. Kieser, John Innes Centre, (Redenbach *et al.*, 1996)

Table 2.2, continued

Plasmid	**Features**	**Antibiotic resistance**	**Reference and/or source**
SCC117	Supercos 1 derivative containing 37.2 kb of *S. coelicolor* chromosomal DNA, including the *sco2529* gene	kanamycin carbenicillin	Gift from H. Kieser, John Innes Centre, (Redenbach *et al.*, 1996)
SCE25	Supercos 1 derivative containing 45.5 kb of *S. coelicolor* chromosomal DNA, including the *prsI*, *arsI* and *sigI* genes	kanamycin carbenicillin	Gift from H. Kieser, John Innes Centre, (Redenbach *et al.*, 1996)
SCE25.2.E09_040 22803W3.seq	SCE25 with an *egfp*-containing transposon inserted in *arsI*	kanamycin carbenicillin	Gift from A Gehring and P. Dyson, (Gehring *et al.*, 2000)
pBluescript	multicopy cloning vector	ampicillin	Stratagene
pCR®2.1 TOPO	High copy number phagemid for direct cloning of PCR products	ampicillin	Invitrogen
pIJ2925	pUC18 derivative with BglII sites flanking the multiple cloning site	ampicillin	(Janssen and Bibb, 1993)
pIJ773	template plasmid for amplification of the apramycin resistance cassette, *aac(3)IV+oriT*	ampicillin apramycin	Bertolt Gust; Plant Bioscience Ltd., Norwich (Gust *et al.*, 2003)
pIJ790	λ-RED (*gam, bet, exo*), *cat araC*, repA101ts	chloramphenicol	Bertolt Gust; Plant Bioscience Ltd., Norwich (Gust *et al.*, 2003)
pUC119	High copy number cloning vector with blue/white selection	ampicillin	(Vieira and Messing, 1987), J. Vieira, Waksman institute of Microbiology, Piscataway, NJ
pUZ8002	RK2 derivative with defective *oriTI (aph)*, driver plasmid for conjugation of *oriT*-containing plasmids from *E. coli* to *Streptomyces*	kanamycin	M. Buttner, John Innes Centre

Table 2.3 Plasmids generated in this study

Plasmid name	Parent	Insert details	Purpose	Antibiotic resistance
pAU240	pCR®2.1 TOPO	Insert is the undigested *Eco*RI-*Bam*HI PCR product CGA2-CGA1, which contains the first 146 bp of *bdtA* and 1.2 kb flanking sequence	*bdtA* in-frame deletion	ampicillin
pAU241	pCR®2.1 TOPO	Insert is the undigested *Bam*HI-*Xba*I PCR product CGA3-CGA4, which contains the final 21 bp of *bdtA* and 1.1 kb flanking sequence	*bdtA* in-frame deletion	ampicillin
pAU242	pUC119	*Eco*RI/*Bam*HI fragment from pAU240 and *Bam*HI/*Xba*I fragment from pAU241, cloned into *Eco*RI/*Xba*I digested pUC119	*bdtA* in-frame deletion	ampicillin
pAU243	pIJ2925	1.3 kb *Eco*RI/*Bam*HI fragment from pAU181, containing the entire *bldD* coding region	Construction of the vector for use in microarray experiments	ampicillin
pAU244	pIJ6902	0.8 kb *Nde*I/*Bgl*II fragment from pAU243, containing the entire *bldD* coding region; *bldD* expression is under control of the promoter *ptipA*	Vector for use in microarray experiments	apramycin thiostrepton
pAU245	pSETΩ	Insert is the PCR product CGA21-CGA50, inserted *Eco*RI/blunt into *Eco*RI/*Eco*RV digested vector: contains the entire *arsI* coding region and upstream intergenic region	Complementation of *arsI* deletion mutant	spectinomycin
pAU246	pHJL400	*Eco*RI/*Xba*I fragment from pAU242: contains the *bdtA* coding region with an in-frame deletion, and 1.1-1.2 kb flanking sequence on either side	*bdtA* in-frame deletion	ampicillin, thiostrepton
pAU247	pCR®2.1 TOPO	Insert is the PCR product CGA64-CGA65 that contains the entire *prsI* coding region with an *Nde*I site engineered into the translation start codon	Construction of *prsI* overexpression vector	ampicillin
pAU248	pCR®2.1 TOPO	Vector contains the PCR product CGA20-CGA67 that contains the entire intergenic region between *prsI* and *arsI*, with 138 bp of *prsI* coding sequence	*prsI* S1 mapping probe	ampicillin

Table 2.3 continued

Plasmid name	Parent	Insert details	Purpose	Antibiotic resistance
pAU249	pIJ2925	Vector contains the *Eco*RI fragment from pAU247; insert includes the *prsI* coding sequence with an *Nde*I site engineered into the translation start codon. Insert orientation minimizes the amount of MCS sequence included with the *prsI* fragment in an *Nde*I/*Bgl*II digest.	Construction of *prsI* overexpression vector	ampicillin
pAU250	pBluescript	1.4 kb *Bam*HI fragment from cosmid 2SCK8, containing the entire *sigQ* coding region and flanking sequence	Complementation of *sigQ* deletion mutant	ampicillin
pAU251	pSETΩ	Insert is the *Bam*HI fragment from pAU250, containing the entire *sigQ* coding region and flanking sequence	Complementation of *sigQ* deletion mutant	spectinomycin
pAU252	pIJ2925	Vector contains the *Eco*RI fragment from pAU247; insert includes the *prsI* coding sequence with an *Nde*I site engineered into the translation start codon. Orientation is the reverse of pAU249.	Construction of *prsI* overexpression vector	ampicillin
pAU253	pCR®2.1 TOPO	Insert is the PCR product CGA21-CGA65, with an *Eco*RI site from CGA21: includes the *arsI*/*prsI* operon, including all intergenic sequence upstream of *arsI*	Second copy of *prsI* for deletion of the chromosomal copy	ampicillin
pAU254	pCR®2.1 TOPO	Insert is the undigested *Bam*HI-*Bgl*II PCR product CGA20-CFU1, containing the *prsI* promoter from -270 to +47 relative to *prsIp1*.	Construction of *prsI-egfp* transcriptional fusion vector	ampicillin
pAU255	pCR®2.1 TOPO	Insert is the undigeested *Bgl*II-*Bgl*II PCR product CFU4-CFU5, containing the *bldD* promoter from -339 to +38 relative to *pbldD*.	Construction of *bldD-egfp* transcriptional fusion vector	ampicillin
pAU256	pSETΩ	Insert is the *Eco*RI fragment from pAU253, from *Eco*RI sites in CGA21 and in pCR®2.1 TOPO	Second copy of *prsI* for deletion of the chromosomal copy	spectinomycin
pAU257	pIJ8660	CFU4-CFU5 transferred from pAU255 through digestion of the pCR®2.1 TOPO *Bam*HI/*Xba*I sites: contains the *bldD* promoter and 51 bp sequence from pCR®2.1 TOPO downstream of the *bldD* promoter	*bldD-egfp* transcriptional fusion vector	apramycin

Table 2.3 continued

Plasmid name	Parent	Insert details	Purpose	Antibiotic resistance
pAU258	pIJ8660	CGA20-CFU1 transferred from pAU254 through digestion of the pCR®2.1 TOPO *Bam*HI/*Xba*I sites: contains the *prsI* promoter and 51 bp sequence from pCR®2.1 TOPO downstream of the *prsI* promoter	*prsI-egfp* transcriptional fusion vector	apramycin
pAU259	pIJ6902	Insert is the *Bgl*II/*Nde*I fragment from pAU249: contains the *prsI* coding region under the control of the promoter *ptipA*	*prsI* overexpression vector	apramycin
pAU260	pCR®2.1 TOPO	Insert is the undigested *Bgl*II-*Bgl*II PCR product CFU3-CFU2, containing sequence from 43 bp inside *sigI* to 4 bp inside *arsI*	Construction of *arsI-egfp* transcriptional fusion vector (first attempt)	ampicillin
pAU261	pIJ8660	CFU3-CFU2 transferred from pAU259 through digestion of the pCR®2.1 TOPO *Bam*HI/*Xba*I sites: contains the *arsI* promoter and 51 bp sequence from pCR®2.1 TOPO downstream of the *arsI* promoter	*arsI-egfp* transcriptional fusion vector (first attempt)	apramycin
pAU262	pCR®2.1 TOPO	Insert is the PCR product CFU3-CGA19, with a *Bgl*II site in CFU3, containing sequence from 43 bp inside *sigI* to 57 bp inside *arsI*	Construction of *arsI-egfp* transcriptional fusion vector (second attempt)	ampicillin
pAU263	pIJ8660	CFU3-CGA19 transferred from pAU262 through digestion of the pCR®2.1 TOPO *Bam*HI/*Xba*I sites: contains the *arsI* promoter and 51 bp sequence from pCR®2.1 TOPO downstream of the *arsI* promoter	*arsI-egfp* transcriptional fustion vector (second attempt)	apramycin

(20% v/v final concentration), flash frozen in a dry ice-ethanol bath, and stored at -80°C.

2.1.4 *S. coelicolor* growth conditions

Standard growth medium for propagation of *S. coelicolor* strains was R2YE agar or liquid medium (Hopwood *et al.*, 1985). For ease of harvesting, solid cultures were grown on cellophane discs (75 mm, 325P discs; Courtalds Films). Cultures were grown at 28-30°C in the dark. Tiger Milk (Kieser *et al.*, 2000) was included when growing auxotrophic strains. Other solid media used were MS agar (20 g/L soy flour, 20 g/L mannitol, 20 g/L agar) (Kieser *et al.*, 2000), minimal medium (Hopwood *et al.*, 1985) with 1% glucose as the carbon source and L-asparagine as the nitrogen source, and Difco Nutrient Agar (DNA, 23 g/L). Liquid cultures were also grown in 2xYT broth (Sambrook *et al.*, 1989), NMMP (Kieser *et al.*, 2000), Super-YEME (0.3% yeast extract, 0.5% peptone, 0.3% malt extract, 1% glucose, 34% sucrose supplemented with 5 mM $MgCl_2$, 0.5% glycine and Tiger Milk) (Hopwood *et al.*, 1985) and Difco Trypticase Soy Broth (TSB, 30 g/L). Antibiotics were added as indicated: apramycin (Provel) to 50 μg/mL, nalidixic acid (Sigma) to 25 μg/mL, spectinomycin (Sigma) to 0.1-1 mg/mL or thiostrepton (gift from S. Lucania at Squibb and Sons, Inc., Institute for Medical Research) to 30-50 μg/mL.

2.1.5 *S. coelicolor* spore and mycelial stocks

S. coelicolor spore stocks were made by growing sporulating strains for several days on R2YE agar on cellophane, then harvesting into Universal tubes containing sterile water. Culture clumps were disrupted by sonication in a sonicating water bath (Branson Bransonic 220), ten minutes sonicating followed by ten minutes on ice, until the biomass was sufficiently disrupted. The resulting suspension was allowed to settle on ice, then filtered through sterile non-absorbent cotton. Spores were pelleted by centrifugation for 10 min at 2900 rpm and 4°C in an International Centrifuge (refrigerated model PR-J) with a swinging bucket rotor. Spores were then resuspended in an equal volume of 40% glycerol (final concentration 20% w/v) and distributed in aliquots for storage at -20°C.

Cultures for mycelial stocks of non-sporulating (*bld*) strains were grown for no more than two days on cellophane on R2YE, then scraped into sterile glass homogenizers (Kontes) and ground in sterile water to produce an even suspension. The homogenate was then pelleted by centrifugation (as for spore stocks) to concentrate the cell mass, and the pellet was resuspended in an equal volume of 40% w/v glycerol, aliquotted, and stored at -20°C.

Cultures for lyophils could be grown either on plates or in liquid culture. Plates were scraped into sterile water and the biomass was then homogenized to break up the mycelia, and washed with sterile 10.3% w/v sucrose. Broth cultures were simply pelleted by centrifugation. In both cases, the resulting mass was suspended in an equal volume of sterile 20% w/v skim milk to make a thick suspension. A few drops of suspension were aliquotted aseptically into drawn lyophilization tubes then flash frozen in a dry ice-ethanol bath before being dried overnight under vacuum (Freezemobile 24, Virtis). Tubes were then flame-sealed and stored at -20°C.

2.1.6 Preparation of competent *E. coli* cells

Calcium competent cells were prepared by the method of Sambrook *et al.* (1989). Aliquots of 50 µL or 100 µL were flash frozen in a dry ice-ethanol bath and stored at -70°C.

Electrocompetent cells were prepared by the method of Gust *et al.* (2003). The cell suspension was divided into 50 µL aliquots and either used immediately or flash frozen in a dry ice-ethanol bath before being stored at -70°C.

2.1.8 Generation of *S. coelicolor* protoplasts

Protoplasts were prepared by the method of Hopwood *et al.* (1985), except that cultures were grown in TSB or R2YE supplemented with 0.5% glycine.

2.2 Transfer of DNA

2.2.1 Transformation of competent *E. coli* cells

Lab-generated, chemically competent *E. coli* cells were transformed as previously described (Sambrook *et al.*, 1989) and those from a commercial source were transformed according to the manufacturer's recommendations. For electroporation, thawed cells and DNA were transferred to a sterile, ice-cold 0.2 cm electroporation cuvette (Molecular BioProducts). The cuvette was transferred to an electroporator (BioRad GenePulser II) set to 2.5 V, 200 Ω and 25 µF. For both transformation and electroporation, the cells were recovered in fresh medium at 37°C for one hour before inoculation on LB agar with appropriate antibiotics. If blue/white selection was used for identification of clones containing insert, 100 mM IPTG and 40 μg/mL X-gal were added to the plates with the antibiotic.

2.2.2 Conjugation into *S. coelicolor*

The procedure used for conjugation from *E. coli* into *S. coelicolor* was as described in Kieser *et al.* (2000). In short, *E. coli* strain ET12567 containing pUZ8002 and the plasmid of interest (containing an origin of transfer; *oriT*) was grown overnight in LB with the appropriate antibiotics, then diluted 1/100 into the same medium for growth to an OD_{600} of 0.4-0.6. The cells were then pelleted by centrifugation and washed twice in an equal volume of LB (without antibiotics) before being resuspended in 0.1 volume of LB.

For conjugation into wild type *S. coelicolor*, approximately 10^8 spores were added to 500 μL 2xYT broth and heat-shocked at 50°C for ten minutes. For conjugation into *S. coelicolor bld* mutants, 0.5 mL of a frozen mycelial stock or the homogenized growth (in 20% glycerol) from one fresh 48 h plate was used without heat shock.

The prepared *S. coelicolor* and *E. coli* suspensions (0.5 mL each) were then mixed and spun briefly in a clinical centrifuge (International) at 3000 rpm. Most of the supernatant was poured off, and the pellet resuspended in the residual liquid. The resulting mixture was diluted from 10^0 to 10^{-6} with sterile water, and 100 μL of each dilution was used to inoculate MS agar plates containing 10 mM $MgCl_2$. The plates were incubated at 30°C for 16-20 h before overlaying with 1 mL water containing 0.5 mg nalidixic acid and plasmid-specific antibiotic (i.e. apramycin for pIJ6902). Incubation was continued at 30°C until the colonies were approximately 2 mm in diameter. These exconjugants were then picked and inoculated onto MS agar containing nalidixic acid and the plasmid-specific antibiotic. Colonies were most often inoculated next onto R2YE with nalidixic acid and plasmid selection, then cultured or stocked as per the needs of the experiment.

2.2.3 Transformation of *S. coelicolor* protoplasts

Transformation of *S. coelicolor* protoplasts was carried out as per Hopwood *et al.* (1985).

2.3 DNA manipulations

2.3.1 Isolation of plasmid DNA from *E. coli*

Plasmid DNA was isolated from *E. coli* using the alkaline lysis procedure (Sambrook *et al.*, 1989). The optional phenol/chloroform purification was always used. Purified plasmid DNA was typically redissolved in 50 μL TE buffer (10 mM Tris-HCl, pH 7.5, 1 mM EDTA). For high-purity plasmid DNA isolation, the Qiagen QiaFilter kit was used as recommended by the manufacturer.

2.3.2 Isolation of chromosomal DNA from *S. coelicolor*

Chromosomal DNA was isolated from *S. coelicolor* cultures by Procedure 3 from Hopwood *et al.* (1985). Most often, a scaled-down version using 5 mL cultures was used, with the solutions reduced accordingly.

2.3.3 Agarose gel electrophoresis

For the determination of sizes of large pieces of DNA (0.8-10 kb), agarose electrophoresis was used. Agarose gels ranged from 0.8% to 1.25% w/v agarose in 1x TBE (90 mM Tris, 89 mM boric acid, 2.5 mM Na_2EDTA). Loading dye (6x; 0.25% bromophenol blue, 60% sucrose) was added to samples to a final concentration of 1x prior to loading. Gels were typically electrophoresed in 1x TBE buffer at approximately 100 V for an hour. The standard molecular weight marker was a *PstI* digest of lambda phage DNA (Roche). To visualize the bands, gels were stained in ethidium bromide solution then exposed to UV light on a transilluminator.

2.3.4 Purification of DNA from agarose gels

DNA was purified from agarose gels using one of two methods. The first method, the Qiagen QIAquick® PCR purification kit, was used according to the manufacturer's protocol for rapid purification of fragments. The second method, trough purification (Zhen and Swank, 1993), was used for high yield purification of DNA. For this purpose, a 1% agarose gel with sufficient lanes to allow a space of at least two lanes between the marker and the DNA to be purified, was prepared with 1x TAE (40 mM Tris-acetate, 1 mM EDTA). The gel was electrophoresed in 1x TAE containing ethidium bromide under standard conditions, until sufficient separation of the bands had occurred. The gel was visualized on a UV transilluminator, and a trough was cut below the band of interest. The gel was then returned to the buffer tank and the trough was filled with a solution of 30% PEG 8000 and 0.5 μg/mL ethidium bromide in 2x TAE.

The gel was run at 130 V for 2-3 min or until the band ran into the trough (monitored by a hand-held UV source - Mineralight® Lamp UVSL-25, Ultra-Violet Products, Inc.). The PEG solution was removed from the trough, and was extracted with phenol/chloroform and chloroform, and the DNA was precipitated with ethanol and sodium acetate.

2.3.5 Polyacrylamide electrophoresis

Polyacrylamide gel electrophoresis was normally used for fragments smaller than 800 bp. Typically, 5% acrylamide gels (29:1 acrylamide:N,N'-methylene bisacrylamide, BioRad) were prepared in 1x TBE, and were run at 200 V in 1x TBE. Small gels were run for 20-25 min, while large gels were run for 2 h. Molecular Marker V (Roche) was used for fragments under 550 bp, and Molecular Marker III (Roche) was used for larger fragments (up to 800 bp). Loading dye was added as described for agarose gels. Gels were visualized by staining in diluted ethidium bromide, then photographing on a UV transilluminator.

2.3.6 Purification of DNA from polyacrylamide gels

Fragments ~0.8 kb and smaller were purified by the "crush and soak" method of Sambrook *et al.* (1989). DNA was electrophoresed on a small 5% polyacrylamide gel, then stained with ethidium bromide and visualized on a UV transilluminator to allow excision of the desired fragment. The gel slice was placed in a 0.6 mL tube that had been punctured with an 18 gauge needle. This tube was in turn placed in a 1.5 mL tube, and the gel forced into the 1.5 mL tube by centrifugation. A volume of elution buffer (0.5 M ammonium acetate, 1 mM EDTA) equal to twice the volume of the crushed gel was added, and the tube was incubated overnight at 37°C in a rotating rack. The gel was sedimented by centrifugation, and as much buffer as possible was transferred to a fresh tube. The crushed gel was then washed at least once with a half-volume of elution buffer. Any residual acrylamide in the combined eluates was sedimented by centrifugation, then the supernatant was transferred to a fresh tube, and precipitated with ethanol and 20 μg glycogen. The purified DNA was washed in 70% ethanol then resuspended in 10 μL milliQH_2O.

2.3.7 Cloning of DNA fragments

PCR fragments were usually cloned into pCR®2.1 TOPO (Invitrogen), using the manufacturer's instructions. Vector and fragment were incubated for 30 min and transformed into One Shot® Top 10 (Invitrogen).

Restriction fragments were transferred from one vector to another using standard restriction enzyme digestion and ligation. For ligation, typical ratios for inster to vector were ~3:1 for sticky ends and ~5:1 for blunt ends. One to three microlitres of the ligated mixture was then transformed into chemically competent or electrocompetent *E. coli* DH5α. Where possible, blue/white selection was used to identify colonies containing a plasmid with insert. Plasmid isolation and restriction enzyme digestion were used to confirm the presence and size of the insert. Colony hybridization (section 2.4.3) was used for large-scale screening.

2.3.8 Polymerase chain reaction

For high-fidelity polymerase chain reaction (PCR) reactions using plasmids as template DNA for the amplification of probes for S1 mapping, DNase I footprinting or electrophoretic mobility shift assay, Expand High Fidelity PCR System (Expand HiFi; Roche) was used. Expand Long Template PCR system (Expand Long; Roche) was often used for amplification directly from chromosomal rather than plasmid DNA. For all other purposes, such as when high fidelity was not required, Taq DNA polymerase (gift of M. Pickard) was used. When chromosomal DNA was used as a template, it was first denatured by room temperature incubation for ten minutes with 0.4 M NaOH and 0.4 mM EDTA, before precipitation with ethanol and redissolution in water. PCR reactions were performed in 20 μL, 50 μL, or 100 μL volumes, containing the following components: 1x reaction buffer (Roche Expand HiFi buffer #2 for Expand HiFi and Taq reactions, Roche Expand Long buffer #3 for Expand Long reactions); 0.2 μM dNTPs; 4 pmol/μL each of forward and reverse primers; DNA template (1-2 μL plasmid or cosmid, or 1 μg denatured chromosomal DNA); and 0.0125 U/μL of Expand enzymes, or varying concentrations of Taq. Reactions were supplemented with 0-6 % DMSO and/or 0-1.5 M betaine to optimize reaction efficiency, as determined for each template and primer pair individually. Reactions were carried out in thin-walled, flat-topped 0.2 mL PCR tubes (Axygen), using either a Perkin Elmer GenAmp PCR System 2400 or a Biometra® T-Personal (Montreal Biotech Inc.) thermocycler. The standard PCR conditions used were: 5 min of denaturation at 95°C; thirty cycles of 30 s at 95°C, 30 s at the annealing temperature, and 60 s at the extension temperature; and, where appropriate (i.e. Topo® cloning), ten minutes at the extension temperature to ensure that the final A residue was added to all fragments. The annealing temperature was separately estimated for each primer pair, and the extension temperature was enzyme-dependent: 72°C for Expand HiFi or Taq, and 68°C for Expand Long. All primers used for amplification of DNA can be found in Table 2.4.

2.3.9 Labelling of DNA

DNA was radiolabelled with ^{32}P by either end-labelling or random primer labelling, depending on the purpose of the experiment and the length of the fragment to be labelled. To end-label DNA [Roche, modified from (Chaconas and van de Sande, 1980)], fragments were incubated at 37°C in kinase reaction buffer (50 mM Tris-HCl, pH9.5, 1 mM spermidine, 10 mM $MgCl_2$, 0.1 mM EDTA, 5 mM DTT) containing 50 μCi $\gamma^{32}P$-ATP (ICN or Amersham) and 1 U kinase (Roche, diluted in 50 mM Tris-HCl, pH7.6, 10 mM DTT, 1 mM EDTA, 50% glycerol) for fifteen or more minutes, then with a further unit of kinase for another fifteen or more minutes. The labelled probe was precipitated with 20 μg glycogen, then redissolved in the appropriate amount of milliQH_2O.

Random primer labelling [Roche, modified from (Feinberg and Vogelstein, 1983)] was used for longer fragments of DNA. The DNA probe and milliQH_2O were mixed (9 μL total) and heated to 90°C for ten minutes to denature the probe. Following cooling on ice, hexanucleotide mix (2 μL; Roche), dNTP mix

Table 2.4 Oligonucleotides used in this study

Name	Sequence *	Relevant Information
BKL37	cgagctggcggacttct	*bldD* internal sequence, +168 to +184 relative to ATG
BKL41	cgccgtcatctacgacc	*bldD* internal sequence, +396 to +412 relative to ATG
BKL50	ccacgacggccttccag	Complementary to *bldD* internal sequence, +102 to 118 relative to ATG
BKL51	gcgcgaattcggcgttcgacgatctcg	Upstream of *bldD*, -400 to -382 relative to *bldD* ATG, with *Eco*RI site: see also CFU4
BKL54	ccgccttcgccaccggt	Complementary to 16S rRNA
BKL104	gcaaaacctcatacagaaaattc	Flanks the polylinker of pIJ6902
BKL105	cacgcggaacgtccgggcttgcac	Flanks the polylinker of pIJ6902
CFU1	gcgcagatctccgggtggtgct*cat*cag	*prsI* reverse primer, complementary to -3 to +15 relative to ATG: sequence is the same as CGA18, with a *Bgl*II site engineered on the 5' end
CFU2	gcgcagatctc*cat*gccgtttctctctc	*arsI* reverse primer, complementary to -14 to +4 relative to ATG: sequence is the same as CGA48, with a *Bgl*II site added on the 5' end, see also CGA62
CFU3	gcgcagatctacgtcgcttggtgggtac	*arsI* forward primer, 25 to 42 bases inside the divergently encoded *sigI* coding region: sequence is the same as CGA63, with a *Bgl*II site added on the 5' end. See also CGA21
CFU4	gcgcagatctggcgttcgacgatctcg	Upstream of *bldD*, -400 to -382 relative to ATG: sequence is the same as BKL51, with a *Bgl*II site engineered on the 5' end instead of the *Eco*RI site
CFU5	gcgcagatctgcggcaggctgtgttgtc	Complementary to *bldD* internal sequence, -43 to -25 relative to ATG: sequence is the same as MAE4, with a *Bgl*II site engineered on the 5' end instead of the *Xba*I site.
CGA1	gcgcggatccgaacgaccgtccgacccg	Complementary to *bdtA* internal sequence, +129 to +146, with *Bam*HI site engineered on the 5' end.
CGA2	gcgcgaattccgcttacggggtgcctcc	1.2 kb upstream of *bdtA*, with *Eco*RI site engineered on the 5' end.
CGA3	gcggatccgtgggggtggaaacggcc	*bdtA* internal sequence, +192 to +208, with a *Bam*HI site engineered on the 5' end.
CGA4	gctctagatccccggcgatgttgaag	Complementary to sequence 1.1 kb downstream of *bdtA*, with an *Xba*I site engineered on the 5' end.
CGA18	ccgggtggtgct*cat*cag	*prsI* reverse primer (overlaps *prsI* start codon), complementary to -3 to +15 relative to ATG: see also CFU1

Table 2.4 continued

Name	Sequence *	Relevant Information
CGA19	ccgtacctcgaccagaag	Complementary to *arsI* internal sequence, +39 to +57 relative to ATG
CGA20	gcgcggatccgccatgcagccggtagtg	*arsI* internal sequence, with *Bam*HI site, 88 to 106 bases upstream of the A of the TGA.
CGA21	gcgcgaattcacgtcgcttggtgggtac	*arsI* forward primer, 25 to 42 bases inside the divergently encoded *sigI* coding region: sequence is the same as CGA63, with an *Eco*RI site added on the 5' end. See also CFU3
CGA40	gggcgcgagatgaacacg	Upstream of *sigQ*, -150 to -133 relative to ATG
CGA40tail	gcgcgaattcgggcgcgagatgaacacg	Upstream of *sigQ*, -150 to -133 relative to ATG: sequence is the same as CGA40, with an *Eco*RI site engineered on the 5' end
CGA41	ttcgcgtcaaccaccgtc	Complementary to *sigQ* internal sequence, +3 to +20 relative to ATG
CGA42	cggaggcgggcgcggacg	Complementary to *prsI* internal sequence, +68 to +85
CGA43	ccctgaaccggctgtgtacg	*prsI* internal sequence, +78 to +59 relative to ATG
CGA45	accgtcggaattccctgaaccaggaagaggccacga*gtg***attccggggatccgtcgacc**	Forward primer for PCR-directed mutagenesis of *sco5718*. Translation start codon indicated by italics.
CGA46	caccaggtgccgcgccgttctctccgaaccggccgg*tca***tgtaggctggagctgcttc**	Reverse primer for PCR-directed mutagenesis of *sco5718*. Translation stop codon indicated by italics.
CGA47	cggtgggaaggcgtagtg	Upstream of *arsI*, -65 to -48 relative to ATG
CGA48	c*cat*gccgtttctctctc	*arsI* reverse primer, complementary to -14 to +4 relative to ATG: see also CFU2 and CGA62
CGA49	gtgttcgtgcgatccggtc	Upstream of *prsI*, -108 to -90 relative to ATG
CGA50	cgtacacagccggttcag	Complementary to *prsI* internal sequence, -75 to -58 relative to ATG
CGA51	tggtggatcggtgaggtg	Upstream of *prsI*, -24 to -14 relative to ATG
CGA52	gaaccccggtgtgcgtgc	Upstream of *sigQ*, complementary to -85 to -68 relative to ATG: reverse of CGA53
CGA53	gcacgcacaccggggttc	Upstream of *sigQ*, -85 to -68 relative to ATG: reverse of CGA52
CGA54	tgcgctcgcggcaccgggcgtcagcacaccgcgtac*atg***attccggggatccgtcgacc**	Forward primer for PCR-directed mutagenesis of *sigQ*. Translation start codon indicated by italics.
CGA55	cccgccccctcgggccggtcccccggcccgtgcct*tca***tgtaggctggagctgcttc**	Reverse primer for PCR-directed mutagenesis of *sigQ*. Translation stop codon indicated by italics.
CGA56	gtgacagtgcggtcactcagtgcgagagtgccggac*atg***attccggggatccgtcgacc**	Forward primer for PCR-directed mutagenesis of *sco2529*. Translation start codon indicated by italics.

Table 2.4 continued

Name	Sequence *	Relevant Information
CGA57	cacctgaatccgcatggacccgtgtctagtacgga *acta***tgtaggctggagctgcttc**	Reverse primer for PCR-directed mutagenesis of *sco2529.* Translation stop codon indicated by italics.
CGA58	cagcaagcagtcacgacctcgggagagagaaa cggc*atg***attccggggatccgtcgacc**	Forward primer for PCR-directed mutagenesis of *arsI*. Translation start codon indicated by italics.
CGA59	gcacacgcacccctccgtcggccggacaaggac ggc*tca***tgtaggctggagctgcttc**	Reverse primer for PCR-directed mutagenesis of *arsI*. Translation stop codon indicated by italics.
CGA62	gcgctctagaggtc*cat*gccgtttctctc	Overlaps *arsI* start codon, complementary to -12 to +7 relative to ATG, has an *Xba*I site: similar to CFU2 and CGA48
CGA63	acgtcgcttggtgggtac	Complementary to *sigI* internal sequence, +26 to +43 relative to ATG: see also CGA21, CFU3
CGA64	cgcgcat*atg*agcaccacccggccctacc	*prsI* forward primer, with engineered *Nde*I site overlapping the *prsI* start codon
CGA65	cctgctggccgaccgagacg	Downstream of *prsI*, complementary to 41 to 60 nt downstream of the TAG
CGA67	cgcgtcctcgaagctcag	Complementary to *prsI* internal sequence, +120 to +138 relative to ATG
DBG14	cgtcaatttcgccaccga	Complementary to *bldG* internal sequence, +53 to +70 relative to GTG
DBG37	gcggtgcagcggatcatc	*tsr* internal sequence, +39 to +56 relative to ATG
DBG38	cgggaagggagaagacgtaac	Complementary to *tsr* internal sequence, +497 to +517 relative to ATG
hrdB-F	aagctgaccagattccggc	*hrdB* internal sequence, +125 to +143 relative to GTG: see also hrdB-S1
hrdB-R	ctctgcggcactgaccatc	Complementary to *hrdB* internal sequence, +207 to +225 relative to GTG
hrdB-S1	gcgcgtatacaagctgaccagattccggc	*hrdB* internal sequence, +125 to +143 relative to GTG: sequence is the same as hrdB-F, with an *Nde*I site added on the 5' end
JST7	cgccagggttttcccagtcacgac	Flanks the pSET152/pSETΩ polylinker
JST8	gagcggataacaatttcacacagga	Flanks the pSET152/pSETΩ polylinker
JWA20	gcgcaagcttgggatcgatcgggtcgg	Upstream of *bldG*, -178 to -194 relative to GTG, with *Hind*III site
KC7	gcgccatatgccagccagatccagaagc	Upstream of *sco1700*, -134 to -117 relative to ATG, with *Nde*I site
KC8	ggcctggagctggtagatg	Complementary to *sco1700* internal sequence, +96 to +114 relative to ATG
KTA-GFP-Rev	ccggtgaacagctcctcg	Complementary to *egfp* internal sequence, +14 to +32 relative to ATG, gift from K. Tahlan
LB5	cgtcagcacaccgcgtac	Upstream of *sigQ*, -18 to -1 relative to ATG
LB6	cggcgctcctggacgtag	Complementary to *sigQ* internal sequence, +89 to +106 relative to ATG

Table 2.4 continued

Name	Sequence *	Relevant Information
M13 forward (-20)	gtaaaacgacggccag	Universal primer, anneals inside the pCR$^{®}$2.1 TOPO sequence flanking the polylinker
M13-reverse	caggaaacagctatgac	Universal primer, anneals inside the pCR$^{®}$2.1 TOPO sequence flanking the polylinker
MAE4	tctagagcggcaggctgtgttgtc	Complementary to sequence upstream of *bldD*, -43 to -25 relative to ATG, with *Xba*I site: see also CFU5
MAE5	ggtaagctttcagagctcgtcgtgggac	Complementary to end of *bldD* internal sequence, +485 to +503 relative to ATG, with *Hind*III site
MAE11	cggtagcaggctcacag	Upstream of *bldD*, -231 to -215 relative to ATG
MAE12	gagctgtttggcgtattcg	Complementary to *bldD* internal sequence, +9 to +19 relative to ATG
MAE16	acgcagcagagtaacgctgcgta	Upstream of *bldD*, -96 to -75 relative to ATG
MAE52	ggttctagagcgctcggaccggtagg	Downstream of *bdtA*, 22 to 38 nt from the TGA, with *Xba*I site
MAE68	cgtagccaggcccgagg	Complementary to *bldN* internal sequence, +26 to +42 relative to GTG
MAE71	gtcaatcgggcacagaagc	Upstream of *bldN*, -166 to -148 relative to GTG
MAE72	gcagaattccgtgtccgtcgccgttcc	Upstream of *bdtA*, -226 to -209 relative to ATG, with *Eco*RI site
MAE73	ggttctagacctaccggtccgagcgctg	Complementary to sequence downstream of *bdtA*, 20 to 38 nt from the TGA, with *Xba*I site
prsIdelF	aaacgtgaactggtggatcggtgaggtgaagcgc tg*atg***attccggggatccgtcgacc**	Forward primer for PCR-directed mutagenesis of *prsI*. Translation start codon indicated by italics.
prsIdelR	ggcggacgggcgggtgcccgcccgagcgccgc aggc*cta***tgtaggctggagctgcttc**	Reverse primer for PCR-directed mutagenesis of *prsI*. Translation stop codon indicated by italics.
sigI-delF	ttcggcagccgtatgccgagaacgatggaggac atc*atg***attccggggatccgtcgacc**	Forward primer for PCR-directed mutagenesis of *sigI*. Translation start codon indicated by italics.
sigI-delR	tcgccccccttcagggcaagcgtcacgcggtgag ga*tca***tgtaggctggagctgcttc**	Reverse primer for PCR-directed mutagenesis of *sigI*. Translation stop codon indicated by italics.

* Restriction sites indicated by underscore
Primer sequence specific to the PCR-targeted mutagenesis cassette indicated in bold.
Italics indicate translation start or stop codon

(3 μL;1.5 mM A, G and T), α^{32}P-dCTP (5 μL; Perkin Elmer), and two units of Klenow enzyme (2 μL; Roche) were added, and the reaction was incubated for one hour at 37°C or overnight at room temperature. The probe was purified by spin column (Micro Biospin® 6, BioRad), and labelling was quantified by Cherenkov counting (in a Beckman LS 3801 scintillation counter) of 1 μL of the labelled DNA.

2.4 DNA analysis

2.4.1 Southern transfer

Southern analysis was carried out according to the protocol of Hopwood *et al.* (1985) using Hybond N (Amersham) nylon membranes. The transfer was allowed to proceed overnight, and then the membrane was then crosslinked under UV (BioRad GS GeneLinker™ UV chamber, program C3). Lanes with marker were cut off for separate hybridization. Membranes were usually hybridized immediately, but if not they were wrapped in plastic wrap and stored at -20°C until needed.

2.4.2 Southern hybridization

Membranes were prehybridized in a glass hybridization tube (Fisher) at a temperature approximately 25°C below the melting temperature of the probe, as determined by the formula $T_m = 81.5°C + 16.6\log M + 0.41\ (\%G+C) - 500/n - 0.61(\%\ \text{formamide})$ where M is the ionic strength (0.45 for 3xSSC) and n is the length of the shortest duplex segment (Hopwood *et al.*, 1985). If an oligonucleotide probe was used, the hybridization temperature was 5°C below the T_d [$T_d = 4(G+C) + 2(A+T)$] (Hopwood *et al.*, 1985). Hybridization solution contained 3x SSC, 4x Denhardt's solution [0.08% w/v Ficoll (MW 400,000), 0.08% w/v bovine serum albumin (Fraction V, Roche), 0.08% polyvinyl pyrrolidone (MW 360,000)], 100μ/mL denatured salmon sperm DNA, and formamide, if necessary to lower the hybridization temperature below 65°C, in a total volume of 10 mL. As stated above, probes were labelled using the random primer method (section 2.3.9). A volume of probe equivalent to two million counts per minute was added to the 10 mL solution and allowed to hybridize overnight. Membranes were washed twice with 2x SSC, 0.1% SDS then once or twice with 0.2x SSC, 0.1% SDS until background radiation on the membrane was reduced to minimal levels. The membrane was then wrapped in plastic wrap and exposed to a phosphor screen (Molecular Dynamics) for visualization of labelled bands. ImageQuant™ software was used for analysis.

2.4.3 Colony hybridization

Colony hybridization was used for large scale screening of colonies when blue/white selection was not possible. Colonies were replica plated on LB agar (for later reference) and LB agar overlayed with a nylon membrane, with the appropriate antibiotics included. Membranes were pre-gridded with an ethanol-proof marker, and the reference plate had a corresponding acetate grid taped to the back. After overnight growth, the reference plate was stored in the fridge and the membrane was subjected to the following treatment. The membrane was soaked for three minutes in 10% SDS, ten minutes in denaturing solution (0.5 M

NaOH, 1 M NaCl), and then twice for five minutes in neutralizing solution (3 M NaCl, 0.5 M Tris-HCl, pH7.5). Finally, the membrane was placed in 2x SSC and the cell debris scrubbed off with a gloved hand. DNA was crosslinked to the membrane using UV (BioRad GS GeneLinker™ UV chamber, program C3) and the membrane was pre-hybridized, hybridized and washed as for a Southern blotting (section 2.4.1) except that the times used were much shorter - one hour for prehybridization and three hours for hybridization. Membranes were wrapped in plastic wrap and exposed to a phosphor screen for visualization of hybridizing colonies.

2.4.4 Sequencing

Sequencing reactions were used to provide ladders for experiments requiring high resolution sizing of products, such as S1 nuclease mapping and DNase I footprinting. For this purpose, sequencing was performed using the Thermo Sequenase Radiolabeled Terminator Cycle Sequencing Kit (USB) and ^{33}P labelled ddNTPs (Amersham), according to the manufacturer's instructions (based on the chain termination method; Sanger et al. 1977). Sequencing reactions contained 0.5 pmol primer, 1x Reaction Buffer (14x concentrate: 260 mM Tris-HCl, pH 9.5, 65 mM $MgCl_2$; USB), 1x dGTP Nucleotide Master Mix (3.5x concentrate: 7.5 μM dATP, dCTP, dGTP, and dTTP; USB), 1 U Thermo Sequenase DNA Polymerase (USB), and 0.5 μL of one of four ^{33}P-labelled ddNTP terminators, in a 7 μL final volume in thin-walled, flat top 0.2 μL tubes (Axygen). Standard PCR cycling conditions, with a 72°C extension temperature, were used as described in section 2.3.8. After completion of the reaction, four microlitres of Stop Solution (95% deionized formamide, 20 mM EDTA pH 8.0, 0.05% bromophenol blue, 0.05% xylene cyanol) was added to each reaction tube. Three microlitres of the terminated sequencing reaction was used for electrophoretic analysis.

Sequencing reactions were typically analyzed on a polyacrylamide gel containing 6% acrylamide (19:1 acrylamide:N,N'-bisacrylamide; BioRad), 8.3 M urea, and 1x TBE. For some DNase I footprinting reactions, a 10% polyacrylamide gel was used. Gels were poured between sequencing gel plates held apart with 0.4 mm spacers, and a twenty-well square-toothed comb was used to create the wells. Gels were electrophoresed at 40W for 1.5-3 h, then transferred to Whatman No. 1 filter paper and dried on a gel dryer (BioRad) at 80°C for two hours (longer for 10% gels). Gels were exposed to a phosphor screen and scanned by either a Molecular Dynamics Model 445 S1 phosphorimager or a Fujifilm FLA-5000. In the former case, the resulting images were analyzed by ImageQuant™ software, and in the latter images were captured by Image Reader FLA-5000 software, processed through ImageGauge V4.22 and examined in Adobe Photoshop.

2.5 RNA purification and analysis

2.5.1 RNA purification and quantification

RNA isolation from *S. coelicolor* was done according to the modified Kirby's protocol described by Hopwood *et al.* (1985). RNA was harvested either from solid-culture time courses grown on cellophane disks on R2YE agar, or from liquid-culture time courses grown in media specific to protocols. Growth times on solid medium were adjusted if growth or differentiation was delayed, but typically mycelia were harvested after 15, 18, 24, 36 and 48 hours of growth. Solid medium-grown cells were scraped directly from the plates into sterile screw cap Universal bottles containing ~2 mL of glass beads and 5 mL modified Kirby mixture (Kieser *et al.*, 2000). Liquid grown cultures were centrifuged quickly (~ 1 min), the supernatant decanted, and the modified Kirby mixture (5 mL) and glass beads (2 mL) added. The resulting mixtures were mixed by vortexing for a total of two minutes (repetitions of 30 s vortexing and 30 s on ice), then 5 mL phenol:chloroform:isoamyl alcohol (25:24:1) was added followed by an additional minute of vortexing. After transfer to a 12 mL polypropylene tube (Sarstedt), the mixture was centrifuged at 8500 rpm, 4°C for 10 min in a Beckman J2-H5 centrifuge using the JA20 rotor. The aqueous layer (containing the nucleic acids) was repeatedly transferred to a fresh tube containing 5 mL phenol:chloroform:isoamyl alcohol and mixed by vortexing, until no interface could be seen. The nucleic acid was precipitated from the aqueous solution with isopropanol and sodium acetate on ice, and either stored overnight at -70°C and then thawed on ice, or processed immediately. Samples were centrifuged for 10 min at 8000 rpm and 4°C, the pellets washed with 95% ethanol, and the nucleic acid redissolved in 450 μL milliQH_2O before the addition of 50 μL 10x DNase digestion buffer (0.5 M Tris-HCl, pH7.8, 0.05 M $MgCl_2$). The RNA solution was transferred to a 1.5 mL tube and digested with two 70 U aliquots of RNase-free DNase (Roche), by incubation for 30 min after each enzyme addition. The reactions were then cleaned up using three phenol/chloroform extractions, two chloroform extractions, and isopropanol precipitation. RNA was collected by centrifugation at 14 000 x g for 10 min, washed with 95% ethanol, and redissolved in 100 μL RNase-free water (Gibco BRL) over either several hours or overnight at 4°C. RNA concentration was quantified by spectrophotometry (A_{260}). Integrity of the RNA and absence of DNA was confirmed by electrophoresis on a 1.2% agarose gel in 1x TBE. Samples were stored as an isopropanol precipitate at -70°C, and the quantification was confirmed and corrected by probing a Northern blot (10 μg RNA) with a probe for 16S rRNA (BKL54) and quantifying the bands with ImageQuant (version 1.2, for Macintosh) software.

2.5.2 Northern analysis

Northern blot analysis was performed as previously described (Williams and Mason, 1985). RNA and DNA (marker) samples were glyoxylated to denature the RNA before separation on an agarose gel. The appropriate amount of RNA (10 μg for a 16S rRNA Northern, otherwise 40 μg) was dissolved in 10 mM sodium phosphate buffer, 50% DMSO and 17% deionized glyoxal to a final volume of 9.5 μL. For the marker (Roche Molecular Weight Marker III or V), 2.5 μL was aliquotted into a tube and the above reagents added. Tubes were incubated at 50°C for one hour, then put on ice before addition of 3 μL of RNA loading dye (50% glycerol, 10 mM $Na_xH_xPO_4$, pH 7.0). Samples were analyzed by electrophoresis on a large 1.25% agarose gel at 58 V in recirculated phosphate buffer (10 mM $Na_xH_xPO_4$, pH 7.0) until the xylene cyanol and bromophenol blue dye fronts were approximately equidistant from each other and the edges of the gel (~4.5 h). RNA was transferred to a nylon membrane (Hybond N, Amersham) as described for a Southern transfer (section 2.4.1). After overnight transfer, the RNA/DNA was crosslinked to the membrane using UV (BioRad GS GeneLinker™ UV chamber, program C3), and the membrane was baked under vacuum at 80°C to remove the glyoxyl adduct from the RNA/DNA. Marker lanes were removed, to be hybridized separately, and the membrane was either hybridized immediately (as in section 2.4.2) or stored in plastic wrap at -20°C. The hybridization probe was prepared by random primer labelling except in the case of BKL54 which was end-labelled (section 2.3.9).

2.5.3 S1 nuclease mapping

High resolution S1 nuclease mapping of mRNA 5' ends was carried out both to locate the transcription start points of previously uncharacterized genes, and to examine the transcriptional profiles of potential BldD targets. The NaTCA method (Kieser *et al.*, 2000) was used for all S1 nuclease protection analyses. DNA probes were designed to cover the potential transcription start point of the gene, often extending some distance into the gene sequence, and containing non-homologous sequence at the 5' end to allow distinction of probe-probe reannealing from full-length protection of the probe. In this way, only the end internal to the gene remained labelled after S1 nuclease digestion, so that expression of divergently expressed genes would not be detected. The probe was most often amplified off a cosmid template, though in some cases it was more convenient to use a cloned PCR product as template to produce a longer non-homologous tail for easier visualization of the full-length protection product.

DNA was labelled using the standard end-labelling procedure (section 2.3.9), then redissolved in RNase-free water (Gibco BRL) to 15 or 30 fmol/μL. Concentration of the probe was either determined by comparison with a known concentration of marker during electrophoresis (before labelling) or by calculation based on the radioactivity of the probe (after labelling). To calculate the activity of the probe in microcuries (μCi) from the number of counts per minute (measured by Cerenkov counting), the following equation was used:

$$\mu\text{Ci(probe)} = \frac{\text{cpm (probe)}}{2.22\text{x}10^6\,\text{dpm}/\mu\text{Ci}}.$$

Complete efficiency is assumed in the conversion between cpm (counts per minute) and dpm (decays per minute); i.e., it is assumed that every decay is counted. Then, the original activity of the probe was determined using:

$$A_0 = \frac{A}{e^{-(0.693/t_{1/2})t}}$$

where A is the activity of the probe counted on a given day (in μCi), A_0 is the activity that would be counted on the day the radioactivity of the isotope was calculated, t is the number of days elapsed since the calibration, and $t_{1/2}$ is the half-life of the isotope (14.3 days for ^{32}P). This then allowed consistent calculation of the amount of DNA in the tube, independent of the age of the isotope. The activity was converted to picomoles of DNA using the following equation:

$$\text{pmol probe} = \frac{\mu Ci\ \text{probe}}{4.5\text{x}10^9\,\mu Ci\,/\,mmol}\,\text{x}\,\frac{10^9\,pmol}{mmol}.$$

The calculation assumed 100% labelling of the probe, but allowed compensation for the age of the isotope so that a consistent amount of probe could be added to the RNA.

Forty micrograms of RNA (stored as an isopropanol precipitate) was sedimented by centrifugation in a microfuge at 13,000 rpm for 10 min, and dissolved in RNase-free water containing 30 fmol of labelled probe. The water was then evaporated in a SpeedVac concentrator (Savant). At the same time, 30 fmol of probe, to act as a control, was added to an empty tube and treated identically. The dried pellets were redissolved in 20 μL sodium-trichloroacetate hybridization buffer (3 M NaTCA, 50 mM PIPES, 5 mM EDTA pH7.0) (Kieser *et al.*, 2000) by repeated heating to 70°C and vortexing. The samples were denatured by heating in a 70°C temperature block for 10-15 min, then quickly transferred to a chiller (Grant LTD 6) set to 45°C and left to hybridize overnight. The following day, 300 μL S1 nuclease digestion buffer [0.28 M NaCl, 30 mM CH_3COONa (pH4.4), 4.5 mM $(CH_3CO_2)_2Zn$, 20 μg partially cleaved denatured calf thymus DNA] containing 400 U S1 nuclease (Roche) was added to each sample, and incubation was performed at 37°C for 45 min. Seventy five microlitres of S1 termination solution (2.5 M ammonium acetate, 0.05 M EDTA) was added to each tube, mixed by vortexing, and the tubes were then placed on ice. The nucleic acid was precipitated with two volumes of ethanol and 40 μg glycogen by resting on ice for at least 30 min, followed by 10 min of centrifugation in a microfuge at maximum speed. Pellets were redissolved in 3 μL formamide loading dye [98% deionized formamide, 10 mM EDTA (pH 8.0), 0.025% bromophenol blue, 0.025% xylene cyanol] by repeated heating to 95°C and vortexing. Samples were denatured for 5 min at 95°C, then analyzed by electrophoresis on a 6% sequencing gel beside a sequencing ladder generated using the internal primer from the S1 probe (section 2.4.4).

2.5.4 Primer extension

Primer extension was used to determine transcription profile differences between wild type and *bldD* mutant *S. coelicolor* strains, and to clarify the transcription start points determined by S1 mapping. Reactions were performed using the *C. therm.* Polymerase Two-Step RT-PCR kit (Roche), and were carried out according to manufacturer's recommendations. The primer of interest was end-labelled with γ^{32}P-ATP and was redissolved in RNase-free water (Gibco BRL) to a concentration of 2.5 pmol/μL. For each time point, 40 μg of RNA was aliquotted and redissolved in 9.76 μL RNase-free or DEPC (diethyl pyrocarbonate)-treated water, then transferred to 0.2 mL thin-walled RNase-free PCR tubes. The following reaction components were added to the tubes in a final volume of 20 μL: 1x RT buffer (Roche), 5 mM DTT, 0.8 mM dNTPs, 17.5 U RNA Guard (Amersham Pharmacia Biotech Inc.), 0.25 pmol/μL ^{32}P-labelled primer, and 6 U *C. therm.* polymerase (Roche). The reaction was incubated at 69°C for one hour, then at 80°C for ten minutes to inactivate the enzyme. Samples were transferred to 1.5 mL tubes and precipitated with two volumes of ethanol, one tenth volume of sodium acetate and 20 μg glycogen. The pelleted DNA was redissolved in 3 μL of formamide loading dye [98% deionized formamide, 10 mM EDTA (pH 8.0), 0.025% bromophenol blue, 0.025% xylene cyanol] by repeated heating to 95°C and vortexing, denatured for five minutes at 95°C, and analyzed on a long 6% sequencing gel as described above (Section 2.4.4). The gel was dried and visualized on a phosphor screen. The screen was scanned on a Molecular Dynamics Model 445 S1 phosphorimager and the data analyzed with ImageQuant™ software.

2.5.5 Microarray analysis

For all cultures grown for microarray work, silanized glassware and springs were used. For microarrays using RNA isolated from sporulating *S. coelicolor* strains, spore stocks were made and titred to ensure consistent, dense growth from a small volume of spores. Spores were germinated for six to eight hours in 20 mL 2x YT in a 250 mL flask containing a stainless steel spring at 30°C and 300 rpm. These cultures were harvested by centrifugation at 2900 rpm for 1 min in an International centrifuge, and the pellets were resuspended in 5 mL NMMP (Section 2.1.4) and sonicated for five seconds at the lowest setting using a 2.5 mm probe (Branson Sonifier 450) to break up clumps. This inoculum was divided between five flasks containing 50 mL NMMP and a sixth flask containing 100 mL NMMP, which received twice the inoculum. The initial OD_{600} was determined, and cultures were grown at 30°C for ten to twelve hours, aiming for an OD_{600} of 0.7-1.0. The six cultures were pooled into a 1 L flask and redistributed to ensure even cell density. The five 50 mL cultures were induced with 30 μg/mL thiostrepton and returned to 30°C with agitation at 300 rpm. The 100 mL culture was kept aside uninduced and harvested immediately by centrifugation. One 50 mL flask was harvested at each of 15, 30, 45, 60 and 120 min post-induction. Immediately following harvesting of each culture, RNA was extracted using the modified Kirby's method (section 2.5.1; Hopwood, 1985). Samples were kept at room temperature until completion of the first phenol extraction to prevent induction of cold shock genes.

For microarrays using *bldD* mutant strains, approximately one millilitre of relatively fresh mycelial stock was grown in 10 mL TSB in a Universal tube containing a stainless steel spring for approximately 36 hours at 30°C with agitation at 300 rpm. Sterile tweezers were used to remove the spring, then the mycelia were harvested by centrifugation at 2900 rpm for 1 min and resuspended in 7 mL NMMP. The inoculum was distributed between six silanized spring flasks with 50 mL or 100 mL NMMP + Tiger Milk. Growth and harvesting were performed as described above.

Synthesis of fluorophore-labelled cDNA and hybridization of the cDNA to the microarrays was performed as described previously (Bignell, 2003; Huang *et al.*, 2005).

2.6 DNA-protein interactions

2.6.1 Electrophoretic Mobility Shift Assay (EMSA)

Electrophoretic mobility shift assays (EMSAs) were used to assess BldD binding to a given DNA fragment. The DNA fragment of interest was end-labelled with γ^{32}P-ATP (section 2.3.9). Binding reactions were performed in 100 mM Tris-HCl, pH7.8, 1.5 M NaCl, 20 mM DTT, 20% glycerol, 1 μg poly d(I-

C)/tube, and 1 ng end-labelled DNA/tube. His$_6$-BldD protein, purified in prior work (Elliot and Leskiw, 1999; Elliot, 2000), was added last, at concentrations varying from 0.5 to 4 μM (10-80 pmol/20 μL). Reactions were mixed gently, and were incubated for 20 min at 30°C in a PCR machine (Perkin Elmer GenAmp PCR System 2400). Samples were loaded on a small, running 8% polyacrylamide gel containing 1.5% glycerol, and were electrophoresed for approximately 20 min at 185 V. Gels were dried, exposed to a phosphor screen, scanned using a Molecular Dynamics Model 445 S1 phosphorimager, and the band intensities quantified using ImageQuant™ software.

The dissociation constant (K_D) was calculated by plotting the intensity of the unshifted (unbound) band against the concentration of BldD added to the reaction. A line of best fit was determined, and the equation of the line used to determine the concentration of BldD at which the unshifted band reached 50% of its original intensity. The reaction can be expressed as $[P]_0$ + DNA $[D] \Leftrightarrow$ DNA-protein complex $[PD]$ where $K_D = [P]_0[D]/[PD]$. At 50% binding, the amount of free and bound DNA are equal, such that the K_D is the concentration of protein added at that point, $[P]_0$. Plotting and calculations were done using Microsoft® Excel.

2.6.2 DNase I footprinting

DNase I footprinting was used to determine the location of the BldD binding site on a given piece of DNA. In order to label the probe DNA specifically on one strand, one primer was end-labelled with γ^{32}P-ATP (section 2.3.9) prior to PCR amplification of the probe. The PCR product was subsequently electrophoresed on a small 5% polyacrylamide gel. The gel was then either exposed for approximately 30-60 s to film or stained with ethidium bromide. In the former case, the region indicating the band of interest was marked clearly on the film and the relevant region of the gel excised. In the latter case, the gel was then visualized with a hand-held UV source and the band excised. Labelled probe was purified by crush and soak as described in section 2.3.6, except that the tube was incubated in a 37°C temp. block, and that the incubation time for smaller fragments was reduced to 1-2 hours.

DNase I footprinting reaction tubes were first set up to contain everything except the protein. The final concentrations of reagents for a 20 μL volume were 200 mM Tris-HCl, pH7.8, 20 mM DTT, 5 mM $MgCl_2$, 2.5 mM $CaCl_2$, and 5 mM KCl. Before addition of the BldD protein, the ^{32}P-labelled probe was redissolved to 0.06 pmol/μL and one microlitre was added to each binding reaction tube. His_6-BldD from prior work (Elliot and Leskiw, 1999; Elliot, 2000) was used at concentrations varying from 0.5 to 4 μM and was added last. Binding reactions were carried out as for the EMSAs (section 2.6.1), then 0.05 U of DNase I (Roche) was added to each tube with gentle mixing. After ten seconds, 50 μL neutral phenol was added to each reaction and the tube was mixed by vortexing. Then, 30 μL milliQH_2O was added to each tube and mixed by vortexing. Phases were separated by centrifugation in a microfuge at maximum speed, and the aqueous phase was transferred to a fresh tube containing 50 μL phenol/chloroform and the extraction repeated. DNA in the final aqueous phase was precipitated with ethanol, sodium acetate and glycogen and redissolved directly into 3 µL formamide loading dye (98% deionized formamide, 10 mM EDTA, pH 8.0, 0.025% bromophenol blue, 0.025% xylene cyanol) by repeated heating to 95°C and vortexing. Samples were denatured by a final incubation for five minutes at 95°C then loaded on a sequencing gel (6% or 10% polyacrylamide, depending on the fragment size). A sequencing ladder, generated using the same primer that was labelled for probe preparation, and purified (unlabelled) probe as the template, was electrophoresed beside the samples to allow identification of the footprinted sequence.

2.6.3 Chromatin immunoprecipitation (ChIP)

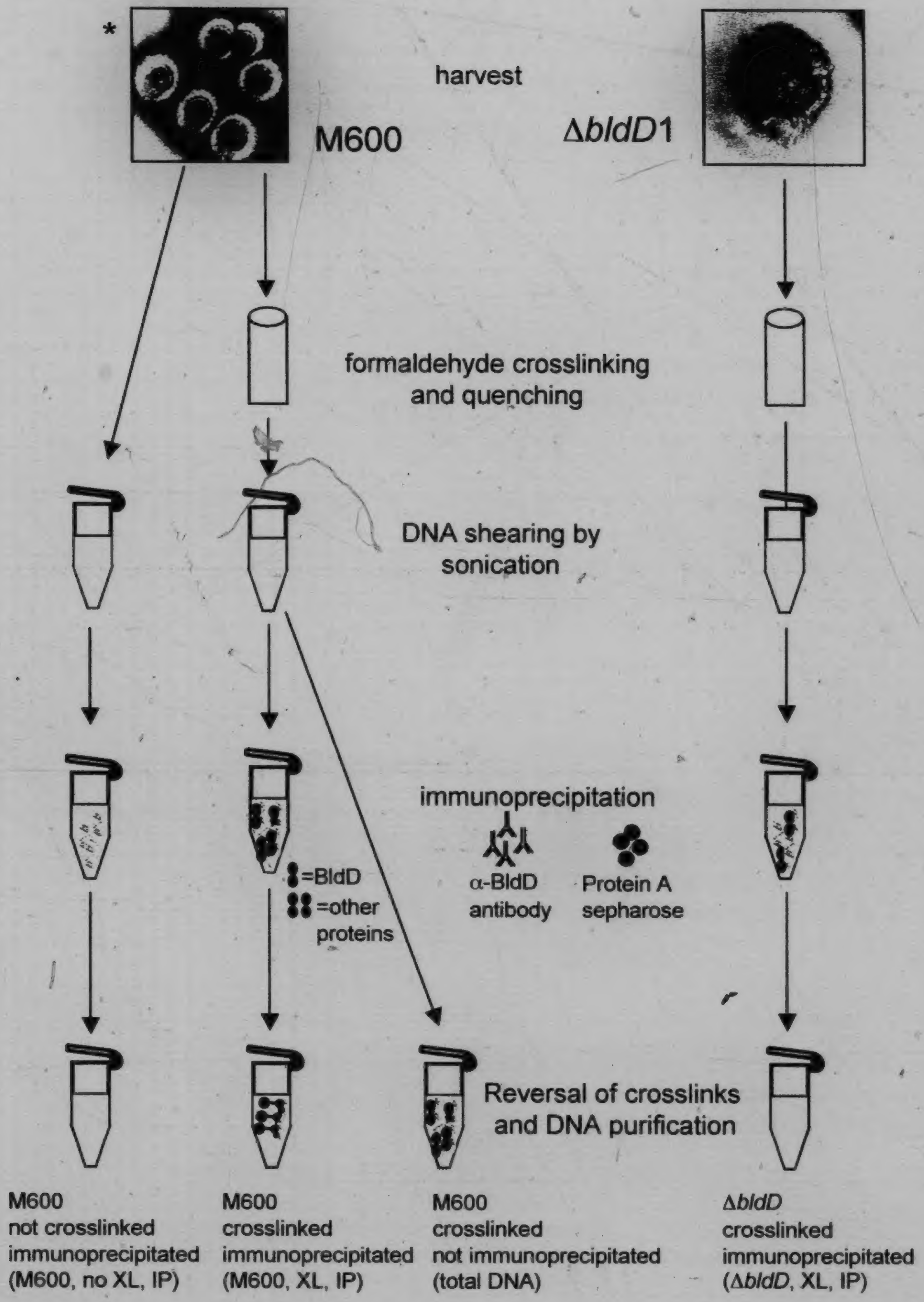
The protocol used for chromatin immunoprecipitation (ChIP) assays was modified by L. Bui from previously described methods (Solomon and Varshavsky, 1985) (Jakimowicz *et al.*, 2002; Shin and Groisman, 2005). Affinity-purified antibodies were used (purification carried out by B. Leskiw and L. Bui). The protocol for affinity purification is found in other work from this lab (Bui, 2006). Two *S. coelicolor* strains were used for the ChIP experiments (summarized in Fig. 2.1): the *bldD* mutant Δ*bldD1*, and its congenic parent M600. Cultures were grown in a 3:2 mixture of Super-YEME (Hopwood *et al.*, 1985): Trypticase Soy Broth (Difco) or on solid R2YE medium with cellophane disks. Cultures were grown for 18, 24, 30 and 48 h for liquid cultures, and for 18, 24 and 48 h for solid cultures. Mycelia were harvested from liquid cultures by centrifugation, then washed three times in 10 mL HEPES solution (50 mM HEPES, pH 7.5, 100 mM NaCl, 1 mM EDTA) and resuspended in 10 mL crosslinking solution (50 mM HEPES pH7.5, 100 mM NaCl, 1 mM EDTA and 1.5% formaldehyde). To harvest

Figure 2.1

Generation of the four ChIP samples. For simplicity, the diagram shows the procedure for solid culture harvest. For liquid culture harvest, mycelia were washed several times following harvesting but prior to crosslinking.

Cell mass (M600 wild type or Δ*bldD*1 mutant) was scraped off plates, and a weighed amount of each was subjected to formaldehyde crosslinking; a sample of M600 was kept aside without crosslinking. The crosslinking was quenched after incubation by the addition of excess glycine. Crosslinked cell mass was washed to remove the crosslinking buffer, then resuspended in sonication buffer. All samples then underwent sonication, centrifugation to clarify the extract, and a second round of sonication to further shear the DNA. One sample was set aside (crosslinked but not immunoprecipitated). The protein in the samples was quantified, then a volume containing 1 mg of protein from each sample was subjected to immunoprecipitation using affinity purified BldD-specific antibodies followed by binding to Protein A sepharose. All samples were subjected to Proteinase K digestion followed by incubation at 65°C for reversal of crosslinks. DNA was purified by phenol:chloroform and chloroform extraction, then precipitation. Samples are labelled at the bottom of the figure; the designation in brackets was used to label gels in later figures.

*Photo courtesy of The John Innes Centre (http://www.jic.bbsrc.ac.uk/science/molmicro/Strept.html/).


*
harvest
M600
Δ*bldD*1
formaldehyde crosslinking and quenching
DNA shearing by sonication
immunoprecipitation
=BldD
=other proteins
α-BldD antibody
Protein A sepharose
Reversal of crosslinks and DNA purification
M600 not crosslinked immunoprecipitated (M600, no XL, IP)
M600 crosslinked immunoprecipitated (M600, XL, IP)
M600 crosslinked not immunoprecipitated (total DNA)
Δ*bldD* crosslinked immunoprecipitated (Δ*bldD*, XL, IP)

mycelia from solid cultures, a measured weight of cell mass (typically 0.5 - 1 g) was scraped into a Universal bottle and mixed with 10 mL crosslinking soution. After harvesting, processing of samples was identical for both culture conditions. One sample of the wild type strain was kept aside in a 1.5 mL tube after harvesting, and was subjected to all of the post-crosslinking treatments as a non-crosslinked, immunoprecipitated control. Formaldehyde crosslinking was allowed to proceed for 15-20 min at room temperature with occasional swirling. The crosslinking was quenched by the addition of glycine to 125 mM and incubation at room temperature for five minutes with swirling, then the cell mass was pelleted by centrifugation and washed twice with ice-cold PBS (Sambrook *et al.*, 1989). Half a gram of crosslinked biomass was transferred to each of two 1.5 mL tubes for sonication. Sonication was performed in PBS with Complete EDTA-free protease inhibitor cocktail (Roche), in 15 s bursts using the microprobe of a Branson Sonifier 450 (constant duty, setting 1). Typically, eight or nine bursts were required to shear most of the DNA to 500-1000 bp fragments; four or five bursts of sonication would be followed by centrifugation at 4°C to pellet the cell debris, with another four or five bursts of sonication to further shear the DNA left in the supernatant (size of sheared fragments verified by agarose gel electrophoresis).

Protein concentration in the cell-free extracts was determined by the BioRad Protein Assay (BioRad; (Bradford, 1976)), using bovine plasma gamma globulin as a standard. For each IP reaction, sonication buffer was added to approximately 1 mg of protein to a final volume of 200 μL, then 500 μL of IP buffer [50 mM Tris-HCl, pH7.5, 300 mM NaCl, 1 mM EDTA, 1 mg/mL bovine serum albumin, 2% NP40 and 1x Complete EDTA-free protease inhibitor cocktail (Roche)] and approximately 5 μg of affinity purified antibody were added. The tubes were incubated overnight on a Labquake® rotisserie at 4°C. At the same time, 150 μL of crosslinked wild type cell extract was set aside and stored at 4°C as a crosslinked total DNA control.

Resin for the precipitation was prepared by swelling 0.1 g Protein A sepharose (Sigma-Aldrich) in 1 mL milliQH_2O at room temperature for 10 min, then sedimenting by centrifugation for two minutes at 3000 x g. The resin was washed twice with IP wash buffer (50 mM Tris-HCl, pH7.5, 300 mM NaCl, 0.5% NP40, 0.1% SDS) and resuspended in one volume of wash buffer to form a 50% slurry.

After overnight incubation of the cell-free extract and antibody at 4°C, 40 μL of the Protein A sepharose slurry was added to each tube. In some experiments, an additional 1.5 mg sheared and denatured salmon sperm DNA was added as a blocking agent. Tubes were then returned to the rotisserie at 4°C for a minimum of 4.5 h. In all following steps, centrifugation of the Protein A sepharose was performed for two minutes at 3000 x g. The bound resin was washed three times with 1 mL wash buffer for five minutes with rotation, three times with 500 μL wash buffer for 30 s with inversion, then twice with 1 mL TE for 30 s with inversion. After removing as much liquid as possible, 100 μL elution buffer (50 mM Tris-HCl, pH8.0, 10 mM EDTA, 1% SDS) was added to the resin and the samples were incubated for 30 min at 65°C. The resin was collected by centrifugation and the eluate was transferred to fresh tubes. A second elution of the resin was with performed using 50 μL elution buffer in the same manner. The eluates of each sample were pooled and 150 μL digestion buffer (50 mM Tris-HCl, pH8.0, 10 mM EDTA, 1% SDS and 400 μg/mL Proteinase K) was added, both to the eluates and to the reserved cell free extract that was to serve as the crosslinked total DNA control (see above), and the tubes were mixed by inversion. The Proteinase K digestion reactions were incubated at 55°C for one hour, then transferred to a 65°C water bath overnight to reverse the formaldehyde crosslinks. The following day, the digests were extracted with phenol/chloroform and chloroform, then precipitated with glycogen, sodium acetate, and ethanol. The DNA from each immunoprecipitation reaction was redissolved in 25 μL TE, and the total precipitated DNA sample (from the reserved cell free extract) was redissolved in 50 μL TE. Samples were analyzed using PCR (section 2.3.8), but using only 20-25 cycles rather than the standard 30. The reduced number of cycles was necessary to minimize the amplification of DNA precipitated because of insufficient shearing. Each gene analyzed, with the primers used, is described in section 3.11

2.7 Gene disruptions

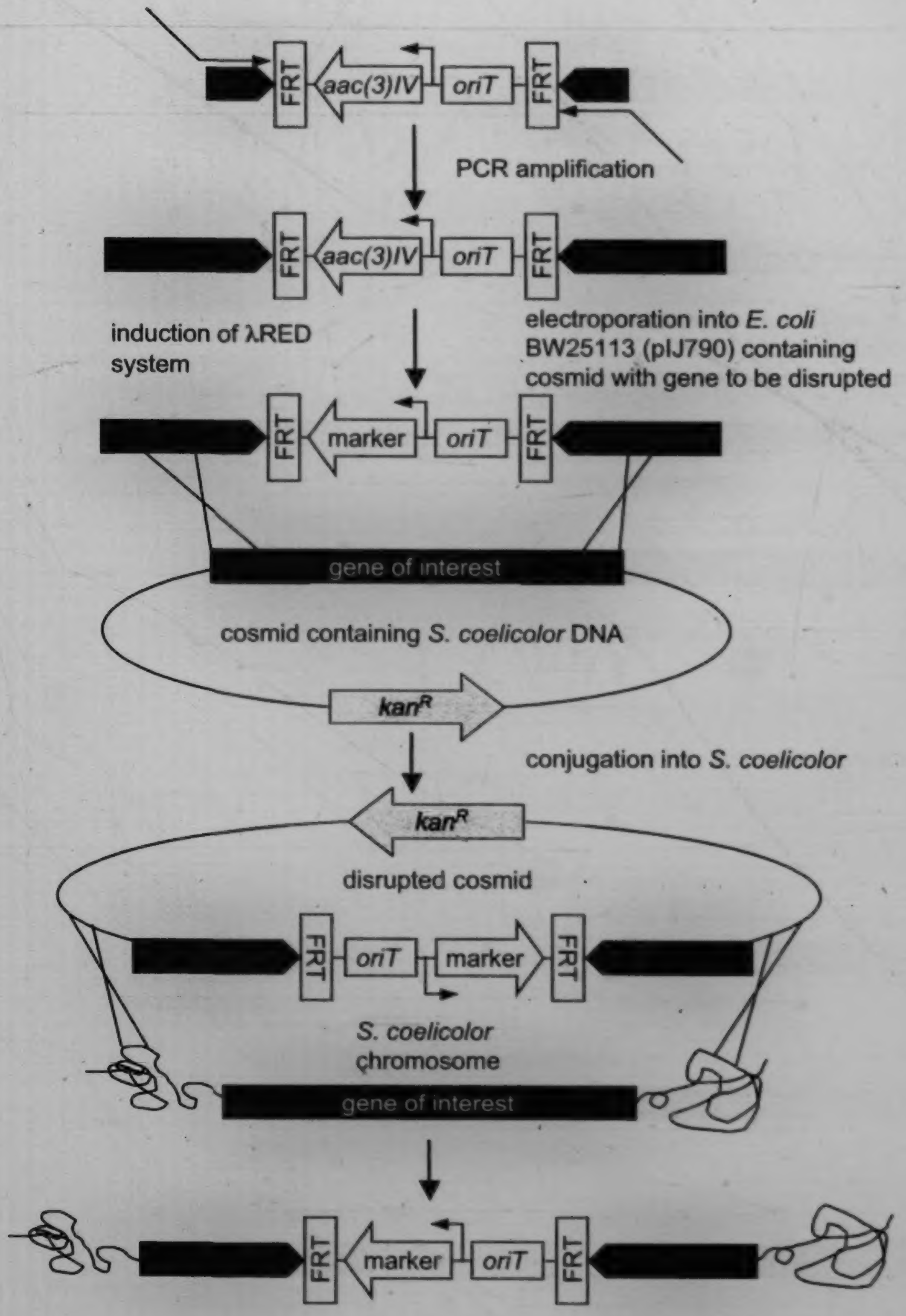
The PCR-targeted gene replacement method of generating marked gene deletions in *S. coelicolor* (Gust *et al.*, 2003) was used to generate all but one of the disruptions in this thesis. The protocol results in a replacement of the entire coding region between the start and stop codons of the gene of interest with a cassette containing an origin of transfer (*oriT*) and an antibiotic resistance marker, in this case the *aac(3)IV* apramycin resistance gene. This procedure is summarized in Fig. 2.2.

Primers used to amplify the disruption cassette had 39 nucleotides matching the *S. coelicolor* sequence adjacent to the gene (including the start or stop codons) at the 5' end, and a sequence matching the disruption cassette (19 or 20 nucleotides) at the 3' end. The 5' sequence was checked using a BLAST-N search (http://www.ncbi.nlm.nih.gov/blast/) for homology to other regions of the *S. coelicolor* genome (more than 30 nt match) to prevent unwanted recombination. The FRT-flanked, *oriT* and *aac(3)IV*-containing cassette to be used as template for PCR amplification of the disruption cassette was removed from the plasmid pIJ773 by digestion with *Eco*RI and *Hin*dIII, and was purified from an agarose gel using the trough method (section 2.3.4). PCR amplification of the disruption cassette was then performed according to the following non-standard conditions: after a two minute denaturation of the template, the product was amplified for ten cycles of 30 s at 95°C, 30 s at 50°C, and 60 s at 72°C, followed by fifteen cycles of 30 s at 95°C, 30 s at 55°C and 60 s at 72°C, with a final extension at 72°C for 5 min. The enzyme used was Expand High Fidelity PCR System (Expand HiFi; Roche); reaction components were as described in section 2.3.8. The product of the PCR reaction (amplified disruption cassette) was purified using the trough purification method described in section 2.3.4.

Preparation of the *E. coli* strain needed for the protocol was done in parallel with the PCR amplification. The cosmid of interest was transferred by electroporation into electrocompetent *E. coli* BW25113 cells already containing

Figure 2.2

Deletion of genes in the chromosome using PCR-targeted mutagenesis. The disruption cassette, containing the *aac(3)IV* apramycin resistance gene and an origin of transfer (oriT) flanked by FRT (FLP recombinase target) sites, was amplified using long primers with 5' ends homologous to the region flanking the gene of interest. The red and blue regions indicate the portion of the primers homologous to the cassette, and the dark grey and purple regions indicate the portion of the primers homologous to the flanking region of the gene of interest. The amplified cassette was then introduced by electroporation into an *E. coli* BW25113 strain already containing pIJ790 (encodes the lambda RED system to promote recombination) and a cosmid (from an *S. coelicolor* genomic library, which contains the gene of interest); the lambda RED system had been induced prior to preparation of the electrocompetent cells. The disrupted cosmid was then transferred from the *E. coli* BW25113 strain by conjugation into a wild type *S. coelicolor* strain (either M145 or M600). Recombinant strains containing the disrupted gene (containing the apramycin resistance cassette) were selected by growth on apramycin only, then screened for the absence of kanamycin resistance (indicating loss of the cosmid). The end result was a strain of *S. coelicolor* with entire coding region of the gene of interest replaced by the antibiotic resistance cassette.


FRT
aac(3)IV
oriT
FRT
PCR amplification
FRT
aac(3)IV
oriT
FRT
induction of λRED system
electroporation into E. coli BW25113 (pIJ790) containing cosmid with gene to be disrupted
FRT
marker
oriT
FRT
gene of interest
cosmid containing S. coelicolor DNA
kan^R
conjugation into S. coelicolor
kan^R
disrupted cosmid
FRT
oriT
marker
FRT
S. coelicolor chromosome
gene of interest
FRT
marker
oriT
FRT

the plasmid pIJ790 (section 2.1.6). The cosmid was from an S. coelicolor genomic library and contained the gene of interest (Table 2.2). pIJ790 has a temperature-sensitive origin of replication and an arabinose-inducible recombination system (λ RED). All strains containing pIJ790 were cultured at 30°C unless specified otherwise, for maintenance of the plasmid.

For the next step, electrocompetent cells were made immediately before use (section 2.1.6). These cultures were grown with ampicillin (100 μg/mL), kanamycin (50 μg/mL) and chloramphenicol (25 μg/mL) to maintain the cosmid and the plasmid, respectively. Arabinose was added during the last three to four hours of growth before preparation of the electrocompetent cells to induce the λ RED recombination system. The amplified disruption cassette was transferred by electroporation into electrocompetent *E. coli* BW25113 cells containing pIJ790 and the cosmid of interest. Cells from this point were incubated at 37°C to promote loss of pIJ790. Cultures were grown on LB agar with ampicillin (100 μg/mL) and kanamycin (50 μg/mL) to maintain the cosmid and apramycin (50 μg/mL) to select for cosmids containing the disruption cassette, and therefore the disrupted gene. To confirm disruption of the gene of interest, cosmids were isolated by the alkaline lysis procedure (section 2.3.1) and analyzed by restriction digestion using an enzyme or pair of enzymes that produced a distinctly different pattern for the recombined and unaltered cosmids.

Disrupted cosmids were transformed into the *E. coli* strain ET12567 containing pUZ8002. From there, the disrupted cosmid was transferred to *S. coelicolor* by conjugation (section 2.2.3). Integration of the disrupted copy of the gene into the *S. coelicolor* chromosome was selected for with apramycin and loss of the cosmid DNA sequences by double recombination to leave the disrupted gene in the *S. coelicolor* chromosome was screened for by loss of kanamycin resistance. Proper recombination of the cassette and replacement of the wild type gene was confirmed by Southern hybridization (section 2.4.2) using three probes: one specific to the gene in question, one to detect the apramycin cassette, and one to detect any gross rearrangements in the chromosome (made by random primer labelling *Sau*3A1-digested cosmid DNA).

2.8 EGFP fluorescence microscopy

Fluorescence microscopy to detect enhanced green fluorescent protein (EGFP) (Sun *et al.*, 1999) was used to determine spatial regulation of gene expression in *S. coelicolor*. The procedure was a modification of the method described previously (Bignell *et al.*, 2005). Fragments presumed to contain intact promoters (containing the transcription start point and surrounding region) were amplified by PCR from *S. coelicolor* cosmids SCE25 or SC9C5 (Redenbach *et al.*, 1996). These fragments were cloned into pCR®2.1 TOPO (Invitrogen), then transferred by *Bam*HI/*Xba*I digestion and subsequent ligation to the *egfp*-containing plasmid pIJ8660 (Sun *et al.*, 1999). The resulting constructs were isolated, transformed into a methylation-deficient strain of *E. coli* (ET12567 with pUZ8002), and conjugated into the wild type *S. coelicolor* strain M145. For each strain, 10 μL spores were inoculated at the base of an ethanol-sterilized No. 0 cover slip (Gold Seal®) inserted at 45° into an R2YE plate. Cultures were grown for two to four days, then the cover slips mounted in mounting medium

(Molecular Probes) on glass microscope slides. M145 containing pIJ8660 was used as a negative control to adjust the settings to minimize detection of autofluorescence, before examining fluorescence in the strains containing pIJ8660 with the promoter regions of the genes of interest. Fluorescence resulting from expression of *egfp* was detected at 505 to 515 nm. A Leica TCS SP2 multiphoton confocal microscope (488 nm argon laser, 100x objective) was used, and imaging was done using the Leica confocal software, version 2.5.

Chapter 3:

Results

Chapter 3. Results

The object of this work was to find new targets of BldD regulation in *Streptomyces coelicolor*. Figure 3.1 shows the *bldD* coding sequence and some of the surrounding region; the fragment of DNA corresponding to the *bldD* locus was used in several experiments, ranging from the induced overexpression for microarray analysis to preparation of probes for DNA binding experiments. The primers and restriction site used for these experiments are indicated.

3.1 Bioinformatic analysis did not reveal new BldD targets

Previous searches for BldD targets had relied on finding *bldD* dependence of genes, then searching for a common sequence in their promoters. Since that type of analysis led to the identification of several target genes, a consensus BldD binding site could be defined - AGTgA (n_m) TCACc (Elliot *et al.*, 2001). The availability of a consensus binding sequence is one factor that made possible a bioinformatics search for additional targets. Equally important to a bioinformatics approach was the publication of the *S. coelicolor* genome sequence (Bentley *et al.*, 2002). Scanning the genome sequence for the BldD binding consensus sequence was a logical approach because it was fast, cheap and relatively easy.

3.1.1 Parameters of the search for BldD binding sites in the *S. coelicolor* genome sequence

A string search (search for a specific sequence, in this case the BldD binding consensus) was performed with the assistance of D. Wishart and

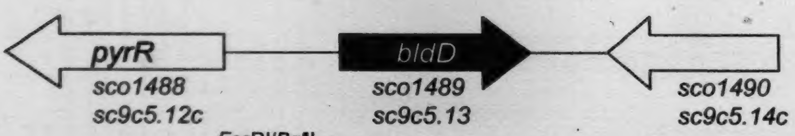
Figure 3.1

Chromosomal gene structure and sequence surrounding *bldD*.

A) Gene structure in the chromosomal area surrounding *bldD*. The genes on the left and right (*sco1488* and *sco1490*) are included for reference. SCO numbers and cosmid designations are included beneath the relevant genes; names are included where possible. Gene sizes and the intergenic regions are drawn to scale.

B) Coding region of *bldD* (bold) with surrounding sequence. The bold sequence at the top is from the divergently encoded gene *pyrR*. Primers used for EMSA, footprinting and EGFP analysis are all marked with an arrow indicating the strand they are derived from; forward primers are marked above the sequence, and reverse primers are indicated below. Wavy lines at the end of some primers indicate engineered restriction sites; the colour of the restriction enzyme is matched to the colour of the primer. The BldD binding site is labelled in red, and the transcription start point is indicated by a bent arrow, also designated p_{bldD}. The *Nde*I site used in creation of the microarray vector, pAU244, is shown in dark blue. The sequence defining BKL37 extends over two lines, so the arrow is coloured green to identify the entire primer.

A



*Eco*RI/*Bgl*II

BKL51/CFU4

B **ccacgtcgtc ggcgcccttg gcgcgttcga cgatctcgtg ggcgatgcgg gtcagcaccc**

gcgcgatgtc gggcccttcg agaacgggcc gggcttcctg ctgctggtcc tgttgcttgt

ccatacgaaa cggacccct tctccgcctc acgggacgga ccttaaagga cgtcggatat

MAE11

gcgccaccta cggtagcagg ctcacagaac ctctccgatg accccctga cactcccccg

tcactcctct gccccacatc gatcacccgc acggagtaat ggatggcgaa taccacggaa

MAE16 p_{bldD}

gagtcggtgc ggaccattcg gcttgacgca gcagagtaac gctgcgtaac ctcacagtg*g*a

gttaccagcc gcgcggccga caacacagcc tgccgcgtcg acaccttgtc cggggagcca

MAE4/CFU5

*Xba*I/*Bgl*II

*Nde*I

tatgtccagc gaatacgcca aacagctcgg ggccaagctc cgggccatcc gcacccagca

MAE12

gggcctttcc ctccacggtg tcgaggagaa gtcccagggc cgctggaagg ccgtcgtggt

BKL50

BKL37

cggttcgtac gagcgcggcg accgtgccgt gaccgtgcag cgcctcgccg agctggcgga

cttctacggc gtccccgtgc aggagctgct gccgggcacc accccgggcg gcgccgccga

gccgccgccg aagctggtcc tggacctgga gcggctggcc accgtgccgg ccgagaaggc

gggcccgctc cagcggtacg cggccacgat ccagtcgcag cgcggtgact acaacggcaa

BKL41

ggtgctctcg atccgccagg acgacctgcg cacactcgcc gtcatctacg accagtcgcc

ctcggtcctc accgagcagc tgatcagctg ggcgtcctg gacgcggacg cgcgccgcgc

ggtggcgtcc cacgacgagc tctgagcccc caccacctca gcagaaacgt gccgccgggg

MAE5

*Hind*III

tggccggaac cgttcgtacg gtcccggcca ccccggcggc ttttacccgt cgtaggctcg

gcgcgcccgt aaggggcgcg gggaactgcg cgaccggcca cggtcggccc gcggccgttc

B. Habibi-Nazhad. The BldD consensus binding sequence used for the genome search was AGTnA (1-20 nt) TCACn, where n was any nucleotide. Because the inverted repeat is not perfect, such that reversing the sequence does not give the same sequence as determining the complementary sequence, this sequence was only searched for in one strand. A perfectly inverted sequence (a palindrome) would be read on both strands in the same location, however any variation in the repeat destroys the palindrome and results in the 5' to 3' sequence on the opposite strand being different than the sequence on the strand being examined. In addition to missing all potential targets in a 5' to 3' direction on the other strand, these search parameters would also have prevented identification of any binding sequences where the two half-sites are not separated by any spacers (such as the first *whiG* binding site - AGTGATCACC) (Elliot *et al.*, 2001), or where the half sites were imperfect matches to the designated consensus, such as the second *bldN* binding site - GGTGA (13 nt) GCACG (Elliot *et al.*, 2001). However, the search was still expected to find many matches within the genome.

3.1.2 Results of the bioinformatics search

From the search of one strand of the genome, 304 potential targets were found. After eliminating sequences found within or downstream of coding regions, 105 consensus matches still remained. Unfortunately, the results of the search were not as useful as hoped. The output format of the program did not give the exact co-ordinates of the sequences, but instead gave the genome co-ordinates of a 60 bp region which contained the putative BldD binding site. Many of the matches found had too great a spacing between half-sites, despite the defined spacing of 1 to 20 nt, or had the half-sites in the wrong order. Two of the known targets were accurately identified - *bldD* and the second BldD binding site upstream of *whiG* - indicating that the search was able to find known BldD binding sites. *bldN* was also found, but should not have been identified by the string search; neither BldD binding site matches the terms of the search. As well, the 60 bp sequence identified upstream of *bldN* by the string search did not contain sequence that matched the search terms. The other known BldD targets, *bdtA* and *sigH*, are encoded on the opposite strand and so were not found by the given search parameters. The more interesting potential targets found by this method are listed in Table 3.1. Targets were found more interesting if they encoded regulators or genes likely to play a role in development.

One surprise from this analysis was the identification of *bldKA* as a potential target. The sequence found by the string search - AGTTA (25) TCACG - was not a particularly good match to the consensus, having too large a spacer between the two half-sites (Table 3.1). This site is found 411 bp upstream of the start codon. There is a good match to the BldD binding consensus closer to the start codon of *bldKA*, although still was somewhat far away from the translational start of the coding sequence - AGTGC (2) TCACG, 206 bp upstream of the start codon. As nothing is known about the transcriptional start of *bldKA*, it is possible that this gene has a very long untranslated region upstream of the coding region, so that the putative BldD binding site might be in an appropriate position for repression of transcription. The nearest gene upstream of *bldKA* is located

Table 3.1 Potential *bldD* target genes identified by bioinformatic analysis of the *S. coelicolor* genome

Cosmid designation	**SCO number**	**Name and function**	**Consensus match**
SC2E1.38	SCO5621	*whiG* - RNA polymerase sigma factor	AGTCA (8) TCACG
SC4H2.32	SCO5811	Transcriptional regulator (2 binding sites)	AGTGA (10) TCACT AGTGA (10) TCACT
SC7E4.13	SCO5216	*sigR* - RNA polymerase sigma factor	AGTGA (14) TCACG
SC9C5.13	SCO1489	*bldD* - DNA-binding protein	AGTAA (12) TCACA
SC9H11.05	SCO7531	Putative AraC-family transcriptional regulator	AGTCA (11) TCACG
SC10F4.35	SCO7662	*cmlR2*, chloramphenicol resistance protein	AGTCA (20) TCACC
SCBAC1C11.02	SCO5999	*sacA* - aconitase	AGTGA (12) TCACT
SCBAC31E11.08	SCO5112	*bldKA* - putative ABC transport system integral membrane protein	AGTTA (25) TCACG
SCC117.02	SCO2529	Putative metalloprotease	AGTCA (6) TCACG
SCC117.05	SCO2532	PhoH-like protein	AGTCA (13) TCACG
SCM11.08	SCO0953	Putative lacI-family transcriptional regulator	AGTCA (6) TCACG

907 bp away, so neither site was in another gene. *bldK* is unlikely to be a BldD target as it is located at a much earlier point than *bldD* in the *S. coelicolor* signalling cascade *bldJ* < *bldK* < *bldA, bldH* < *bldG* < *bldC* < *bldD, bldM* (Chater and Chandra, 2006). Since both *bldD* and *bldK* fit neatly into the cascade it would be difficult to explain how *bldD* could directly regulate the transcription of *bldK*. As further evidence that *bldK* is not likely to be a bona fide BldD target, results from *bldKA* from the microarray analyses that were performed later show no significant change in transcription of *bldKA*, or the rest of the *bldK* gene cluster, upon induction of *bldD*.

As stated above, the parameters used in this search were not expected to identify all of the known BldD binding sequences on the strand examined (for example the known BldD binding sites upstream of *bldN*). Therefore, a second search of the same strand was undertaken, but allowing more variation in the sequence. This second search resulted in thousands of sequences identified, a volume of data too large to manage within the scope of this project. When this was considered together with the results from the first (more limited) search which had indicated that the results were unreliable (many false positives), it became apparent that methods other than simple sequence analysis were necessary in searching for previously uncharacterized targets of BldD regulation.

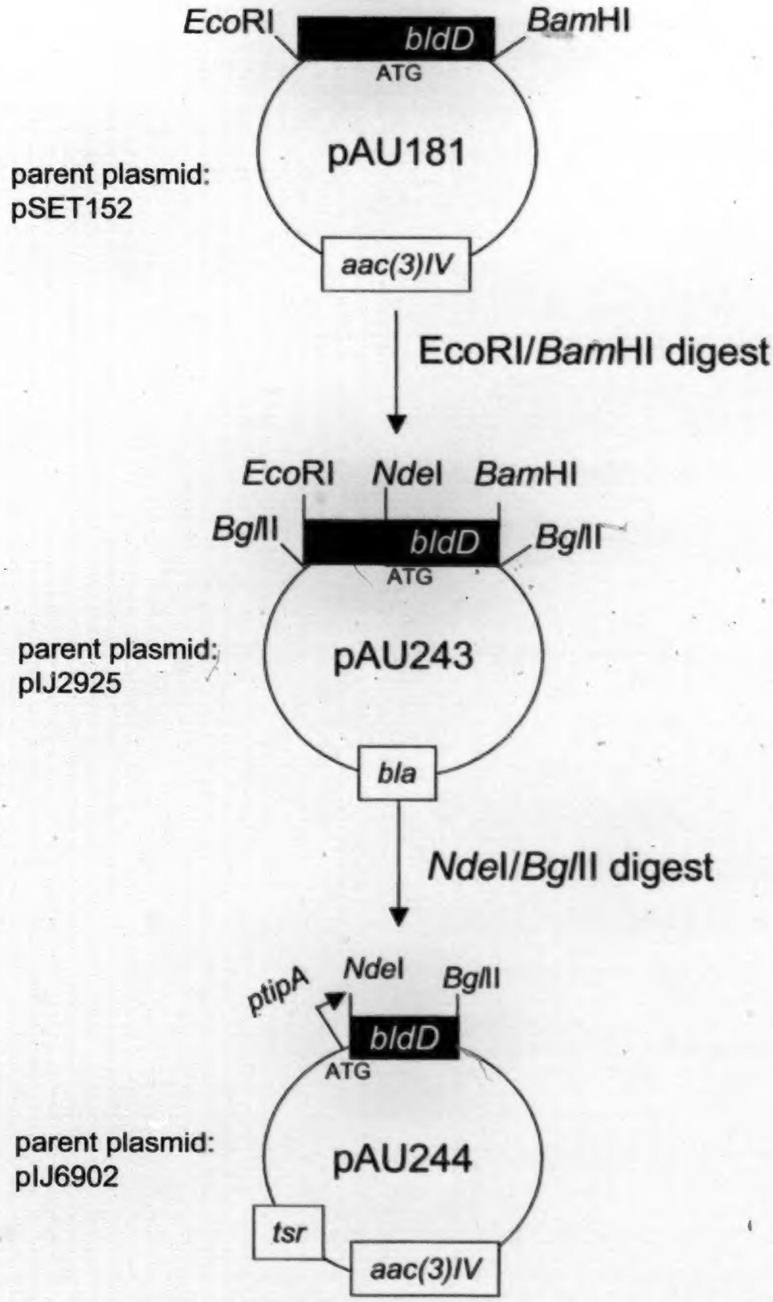
3.2 Microarray analysis revealed new BldD targets

3.2.1 Construction of the vector for controlled overexpression of *bldD*

Given the problems with searching the sequence for BldD binding consensus matches, an alternative method was needed for large-scale scanning of the genome. Microarray analysis allows simultaneous scanning of thousands of genes - in the particular array used, 7071 (90.4%) of the 7825 annotated protein coding genes in the *S. coelicolor* genome (Bentley *et al.*, 2002). The first step in the use of microarray analysis to identify BldD targets was to create a vector that would allow controlled overexpression of *bldD* (Fig. 3.2). The previously described plasmid pAU181 (Elliot *et al.*, 1998) contains the *bldD* coding region on a 1.3 kb *SphI*/*XmnI* fragment, cloned into the *Eco*RV site of pSET152. This plasmid was digested with *Eco*RI and *Bam*HI (from the pSET152 polylinker) and the 1.3 kb fragment ligated into pIJ2925 (Janssen and Bibb, 1993), generating the plasmid pAU243. The polylinker of pIJ2925 is flanked on either side by *Bgl*II sites, allowing inserts to be cloned directionally (from either orientation in pIJ2925) into the integrating overexpression plasmid pIJ6902 (Huang *et al.*, 2005) to place cloned genes under the control of the thiostrepton-inducible promoter *ptipA*. pAU243 was digested with *Nde*I and *Bgl*II and the resulting 0.8 kb fragment ligated into pIJ6902 to make pAU244; the *Nde*I site overlaps the *bldD* start codon (Fig. 3.1), and the 0.8 kb *Nde*I/*Bgl*II fragment contains the entire 504 bp *bldD* coding region. The integrity of the insert was confirmed by sequencing using the primers BKL104 and BKL105 (Table 2.4).

Figure 3.2

Construction of a plasmid vector (pAU244) for inducible expression of *bldD*. Plasmid names are indicated within the plasmids, and the parent plasmids are indicated to the left. pAU181 was digested with *Eco*RI and *Bam*HI (within the pSET152 polylinker) to liberate the 1.3 kb fragment containing the *bldD* coding region. This fragment was ligated into similarly digested pIJ2925, generating pAU243, then digested with *Nde*I and *Bgl*II to liberate a 0.8 kb fragment; the *Nde*I site overlaps the *bldD* start codon (ATG), and the fragment includes the entire *bldD* coding sequence. This fragment was then ligated into pIJ6902 behind the thiostrepton-inducible promoter *ptipA* to create pAU244, an *E. coli* - *Streptomyces* shuttle plasmid with *bldD* expression under the control of *ptipA* and with both thiostrepton and apramycin resistance. *aac(3)IV* is the apramycin resistance gene, *tsr* is the thiostrepton resistance gene, and *bla* is the ampicillin resistance gene.


EcoRI
bldD
BamHI
ATG
pAU181
parent plasmid:
pSET152
aac(3)IV
EcoRI/BamHI digest
EcoRI
NdeI
BamHI
BglII
bldD
BglII
ATG
parent plasmid:
pIJ2925
pAU243
bla
NdeI/BglII digest
ptipA
NdeI
BglII
bldD
ATG
parent plasmid:
pIJ6902
pAU244
tsr
aac(3)IV

3.2.2 Insertion of pAU244 into a wild type strain

pAU244 was transformed into the non-methylating *E. coli* strain ET12567 harbouring the non-transmissible helper plasmid pUZ8002, then introduced into the wild type *S. coelicolor* strain M600 by conjugation. At the same time, the parent plasmid pIJ6902 was separately introduced into the same strain. Integration of the plasmids into the chromosome was confirmed by PCR with the primers BKL104-105, using the purified pAU244 plasmid as a positive control (Fig. 3.3 A).

3.2.3 Microarray analysis of wild type *S. coelicolor* in the presence of *bldD* overexpression

For the pregermination, 20 μL of M600 + pAU244 spores and 40 μL of M600 + pIJ6902 spores were used to inoculate 20 mL 2xYT. This culture was then used as described in section 2.5.5 to inoculate NMMP cultures. Strains were grown in NMMP to an OD_{600} of 0.77 (pAU244) or 0.71 (pIJ6902) before pooling and induction. RNA was harvested at six time points: prior to induction, then at 15, 30, 45, 60 and 120 minutes post-induction. Integrity of the RNA and absence of DNA in the samples was confirmed by agarose gel electrophoresis.

For the microarray hybridization, fluorescently-labelled cDNA was made from the RNA samples; uninduced samples were labelled with Cy5 (red) while induced samples were labelled with Cy3 (green). The uninduced samples were used as the reference to which all other samples were compared. Thus, genes repressed by BldD (transcript levels decreasing as *bldD* is induced) would have

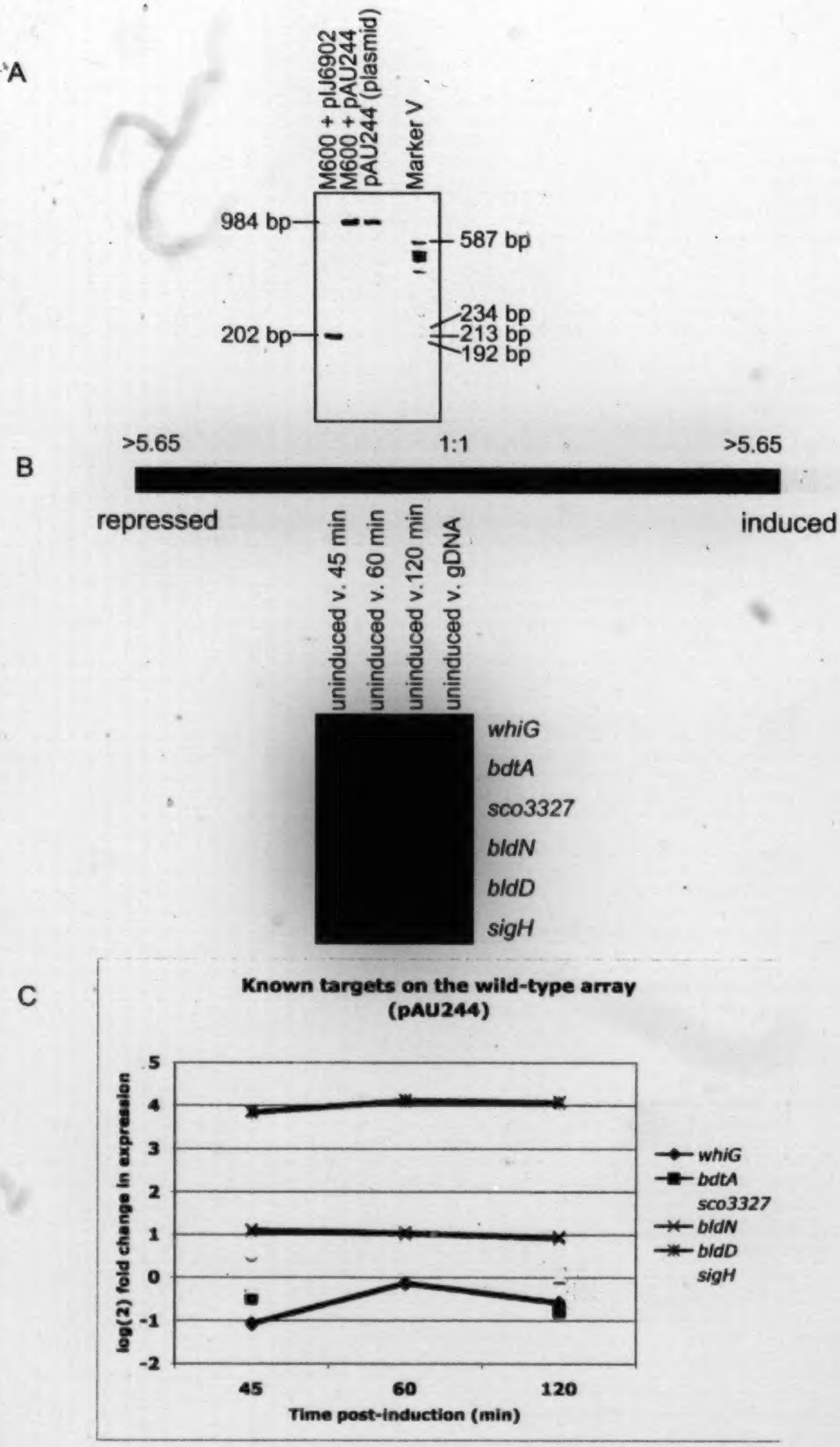
Figure 3.3

Microarray analysis after induced overexpression of *bldD* in wild type *S. coelicolor*.

A) Confirmation by PCR of insertion of pAU244 into the *S. coelicolor* chromosome. Chromosomal DNA was isolated from *S. coelicolor* M600 containing either pAU244 (M600 + pAU244) or the parent plasmid pIJ6902 (M600 + pIJ6902). The primers used were BKL104 and BKL105, which anneal on either side of the multicloning site of pIJ6902. A larger band (984 bp) is amplified when the coding region of *bldD* is inserted, as compared to the band amplified from unmodified plasmid (202 bp). The band amplified from purified plasmid (pAU244) was used as a size marker. The size of selected bands from Molecular Weight Marker V (Roche) is shown on the right.

B) Cluster image of microarray results using RNA harvested from cultures of the wild type *S. coelicolor* strain M600 with integrated pAU244. RNA was harvested before induction, then 45, 60 and 120 minutes after induction of *bldD* expression with thiostrepton. cDNA synthesized from the induced samples (labelled red) was hybridized in competition with cDNA synthesized from the uninduced sample (labelled green). As a control to assess basal transcription levels, the cDNA synthesized from the uninduced RNA sample (labelled red) was hybridized against genomic DNA (labelled green). Grey squares indicate time points for which no data were available. Each gene is identified by name where possible, otherwise by SCO number (the number assigned by the *S. coelicolor* genome project).

C) Plot of $\log_2$ fold change in expression of published BldD targets, in the strain with pAU244. Data from the microarrays came in the form of the $\log_2$ of the ratio of signal from induced samples over signal from uninduced samples. In this way, an increase in expression is seen as a positive number, a decrease in expression is seen as a negative number, and an absence of change gives a value of zero. Lines connect adjacent points; absence of a line indicates lack of data for a particular time point.


A
M600 + pIJ6902
M600 + pAU244
pAU244 (plasmid)
Marker V
984 bp
202 bp
587 bp
234 bp
213 bp
192 bp
B
>5.65
1:1
>5.65
repressed
induced
uninduced v. 45 min
uninduced v. 60 min
uninduced v.120 min
uninduced v. gDNA
whiG
bdtA
sco3327
bldN
bldD
sigH
C
Known targets on the wild-type array (pAU244)
log(2) fold change in expression
5
4
3
2
1
0
-1
-2
45
60
120
Time post-induction (min)
whiG
bdtA
sco3327
bldN
bldD
sigH

an increasingly green signal over time, while genes induced by BldD (transcript levels increasing as *bldD* is induced) would appear red. Use of the control strain allowed removal from the analysis of genes affected by addition of the antibiotic inducer (thiostrepton) or by other factors unrelated to the effects of BldD.

The results were initially surprising. Although *bldD* itself was clearly induced, none of the known targets were affected by induction of *bldD* (Fig. 3.3, B and C). An additional array was used to determine the basal level of *bldD* expression: RNA harvested from uninduced *S. coelicolor* M600 + pAU244 (labelled red) and *S. coelicolor* genomic DNA (labelled green) were hybridized to the same array. There is one copy of each gene per copy of the genome, thus this experiment would demonstrate if a gene was expressed at a significant level (red signal) or silent (or weakly expressed) under the conditions tested (green signal). This experiment demonstrated that basal levels of *bldD* transcript are very high (Fig. 3.3 B). The assumption was made that the protein levels were also therefore very high. The significance to the experiment is in the amount of protein relative to the number of target sites; if all target sites are saturated then additional transcript (and protein) would not affect target gene expression levels. Put differently, if the genes are already maximally repressed, then additional BldD would not make a difference to target transcript levels. Overexpression of *bldD* in a strain without functional BldD protein, however, should result in a visible difference in target gene transcript levels. With this in mind, the experiment was repeated, but with the expression plasmids (pIJ6902 and pAU244) conjugated into the *bldD* mutant strain 1169, which produces no stable BldD protein (Elliot, 2000).

3.2.4 Insertion of pAU244 or pIJ6902 into the *bldD* point mutant (1169)

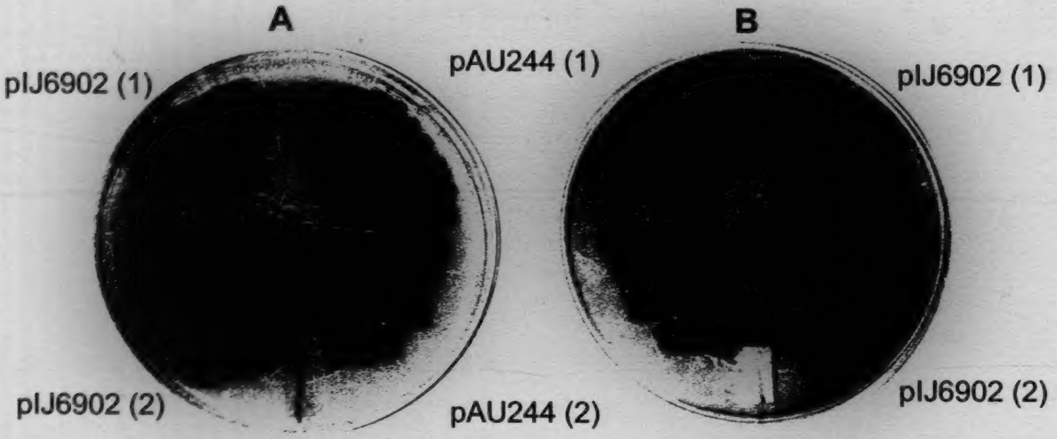
The plasmids pIJ6902 (has thiostrepton inducible *ptipA*) and pAU244 (pIJ6902 with *bldD* under the control of *ptipA*) were transformed into protoplasts of the *bldD* point mutant strain 1169; the *bldD* null mutant, lacking the entire *bldD* coding region (Elliot *et al.*, 2003b), had not yet been made when this part of the project was underway. It had been shown previously (Elliot, 2000) that BldD Y62C, found in strain 1169, is unstable and that there is no BldD protein present in *S. coelicolor* 1169 mutants at any stage of growth. With no active protein present in the absence of induction, any BldD produced by the integrated vector should have an observable effect.

Following transformation of the plasmids into the protoplasts, proper expression and function of the *bldD* gene under control of *ptipA* was confirmed by comparison of the growth of the resulting strains on R2YE with apramycin (Fig. 3.4, A and B) or thiostrepton (Fig. 3.4, C and D). Although pIJ6902 contains *aac(3)IV*, the gene allowing apramycin resistance, growth on this antibiotic was not expected to affect *bldD* transcription and was used as a control for uninduced expression from *ptipA*. Growth on thiostrepton induces expression from *ptipA*; full complementation of the mutant phenotype would indicate that the thiostrepton-induced production of *bldD* results in sufficient levels of active protein for normal growth. As expected, full complementation (i.e. both aerial mycelium and antibiotic production were restored to wild type levels) of

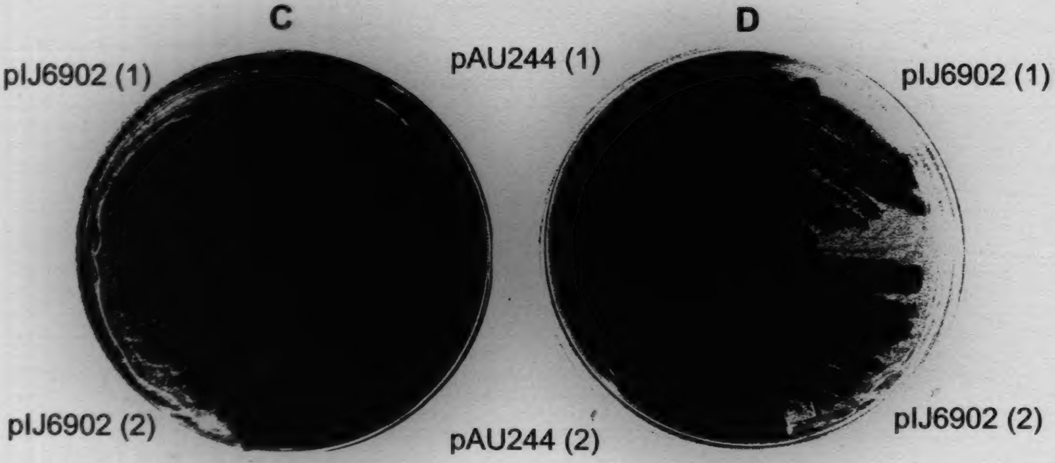
Figure 3.4

The *S. coelicolor bldD* mutant (1169) is complemented by induced expression of *bldD* from plasmid pAU244. *S. coelicolor bldD* mutant strain 1169 with either pIJ6902 or its *bldD*-containing derivative pAU244 was grown on R2YE at 30°C for three days. Plates A and B are top and bottom views, respectively, of the strains grown on R2YE with apramycin; pIJ6902 contains *aac(3)IV*, the gene for apramycin resistance. Plates B and C are top and bottom views, respectively, of the same strains grown on R2YE with thiostrepton, the inducer for overexpression from the *ptipA* promoter of pIJ6902. Plates were inoculated with duplicate strains of each type (labelled 1 and 2).

R2YE + apramycin


A
pIJ6902 (1)
pAU244 (1)
pIJ6902 (2)
pAU244 (2)
B
pIJ6902 (1)
pIJ6902 (2)

R2YE + thiostrepton


C
pIJ6902 (1)
pAU244 (1)
pIJ6902 (2)
pAU244 (2)
D
pIJ6902 (1)
pIJ6902 (2)

1169 + pAU244 was seen on thiostrepton, while apramycin had no effect on the phenotype of the strain.

3.2.5 Microarray analysis of RNA isolated from the *bldD* mutant (1169) *S. coelicolor* after induced *bldD* overexpression

3.2.5.1 Confirmation of repression of known targets by *bldD* induction

The *S. coelicolor* 1169 + pAU244 strain was induced at an OD_{600} of 0.65 and the 1169 + pIJ6902 control strain was induced at an OD_{600} of 0.63. RNA was harvested at six time points; prior to induction, then at 15, 30, 45, 60 and 120 minutes post-induction. Again, integrity of the RNA and absence of any DNA contamination was confirmed by agarose gel electrophoresis. Quantification of the RNA samples was confirmed by Northern analysis using a probe for 16S rRNA (BKL54). Induction of *bldD* transcription was confirmed by Northern analysis using random primer-labelled BKL37-MAE5 as a probe for *bldD* transcript (Fig. 3.5). Transcript is seen even in the absence of inducible *bldD* because the mutation in *S. coelicolor* 1169 is a point mutation that affects the structure of the protein, resulting in deregulation of *bldD* expression.

To ensure that the *bldD* induction would produce the expected repression of known BldD targets, the transcription patterns of both *bldN* and *bldD* were examined in the RNA samples by S1 nuclease protection analysis (Figs. 3.6 and 3.7). The probe used for *bldN* was a DNA fragment, MAE71-68, end-labelled then digested with *Cla*I to produce a 181 bp fragment labelled at only one end (Elliot *et al.*, 2001). The 530 bp probe used for *bldD* (BKL51-50) had a

Figure 3.5

Northern analysis of *bldD* transcripts performed on RNA isolated for microarray analysis. The control strain is 1169 with the integrating microarray plasmid pIJ6902, while the test strain is 1169 containing pAU244 (pIJ6902 with the *bldD* coding region inserted behind the thiostrepton-inducible promoter). Minutes post-induction indicates the time the RNA was harvested after addition of thiostrepton; zero time was harvested without addition of thiostrepton. The probe used for *bldD* was random primer labelled 345 bp DNA fragment, generated using primers BKL37 and MAE5 (Table 2.4), internal to the *bldD* coding region. An oligonucleotide probe for 16S rRNA (BKL54) was used to probe the same membrane as a control for amounts of RNA in each lane.

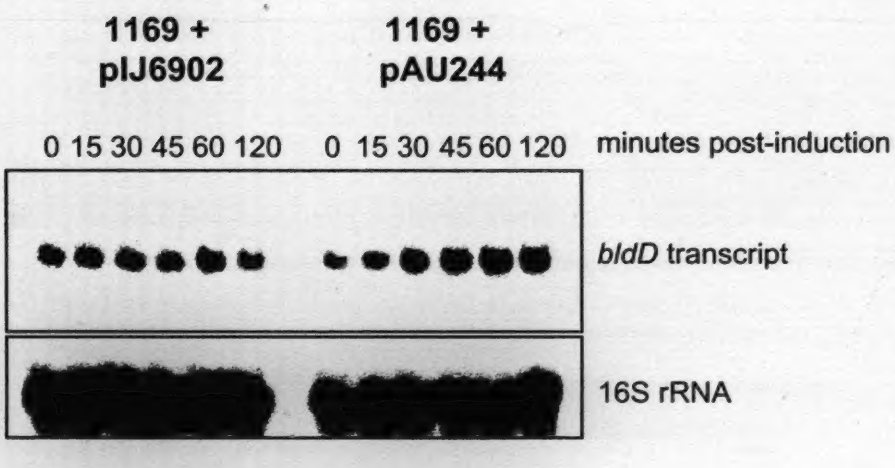

1169 +
pIJ6902
1169 +
pAU244
0 15 30 45 60 120
0 15 30 45 60 120
minutes post-induction
bldD transcript
16S rRNA

Figure 3.6

bldD induction represses *bldN* expression in a *bldD* mutant strain. High resolution S1 nuclease mapping performed on RNA isolated for microarray analysis (from the *bldD* mutant 1169 with pIJ6902 or pAU244). Minutes post-induction indicates the time the RNA was harvested after addition of thiostrepton; zero time was harvested without addition of thiostrepton. The probe for S1 protection of the isolated RNA was a DNA fragment, MAE71-68, end-labelled then digested with *Cla*I to produce a 181 bp fragment labelled at only one end. A sequencing ladder was loaded to the left of the samples to allow determination of the size of the digestion products. P+S1 indicates probe taken through the same S1 nuclease digestion procedure as the samples, but without RNA. The same volumes of RNA were analyzed by Northern hybridization using a probe for 16S rRNA (BKL54) to control for the amount of RNA in each lane.

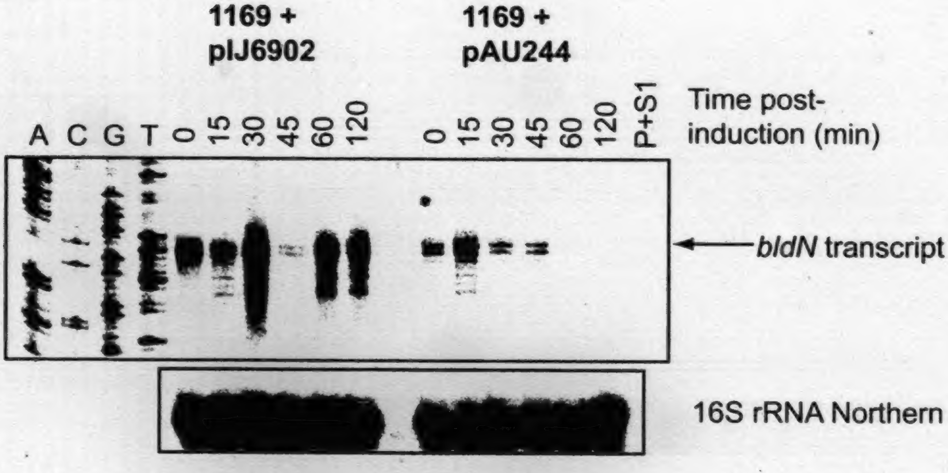
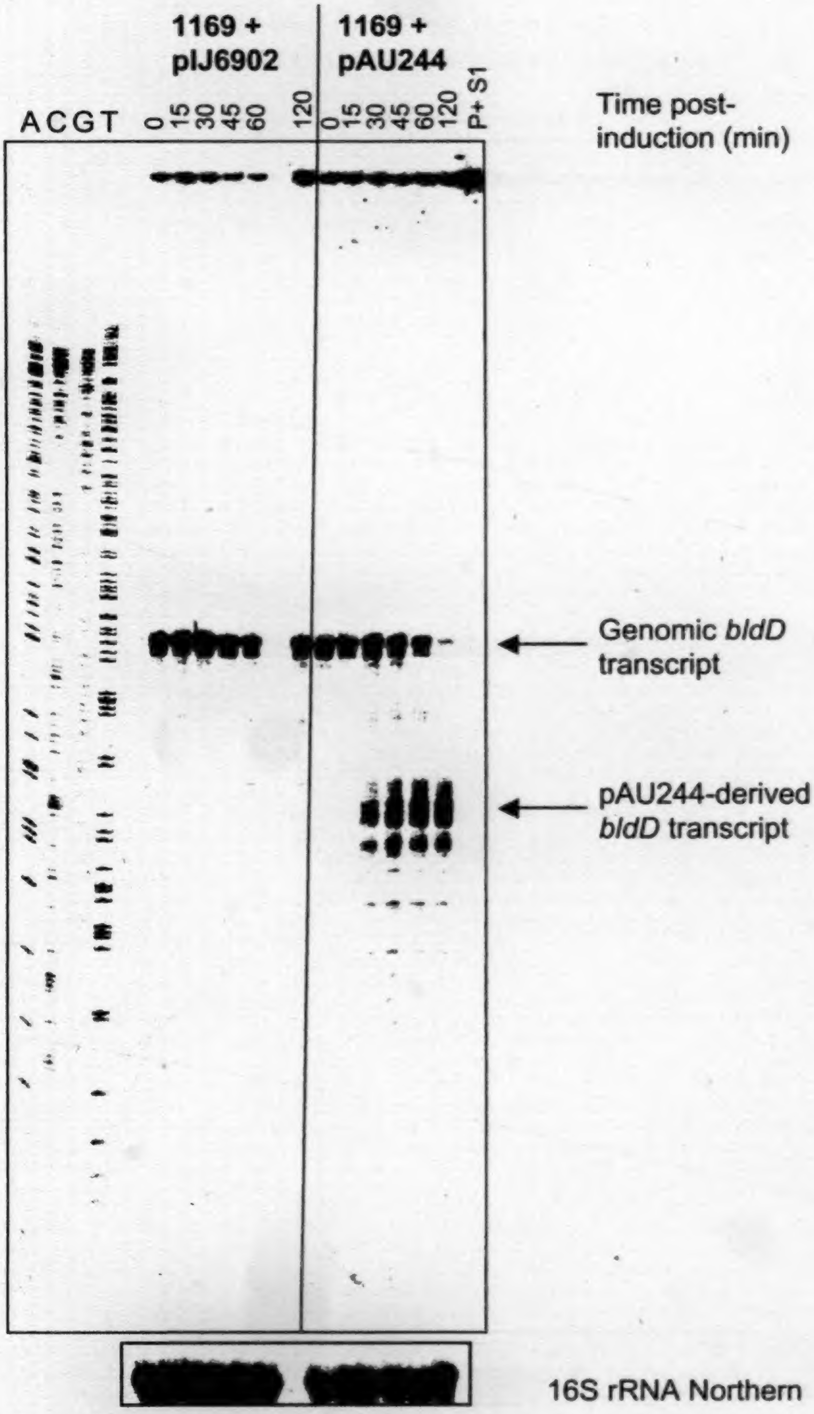

1169 +
pIJ6902
1169 +
pAU244
A
C
G
T
0
15
30
45
60
120
0
15
30
45
60
120
P+S1
Time post-induction (min)
bldN transcript
16S rRNA Northern

Figure 3.7

High resolution S1 nuclease protection mapping of *bldD* transcripts performed on RNA isolated for microarray analysis (from the *bldD* mutant 1169 with pIJ6902 or pAU244). Minutes post-induction indicates the time the RNA was harvested after addition of thiostrepton; zero time was harvested without addition of thiostrepton. The probe for S1 protection of the isolated RNA was the 527 bp PCR product BKL51-50, amplified, purified then end-labelled. Genomic and plasmid-derived *bldD* transcripts are indicated by arrows. P+S1 indicates probe taken through the same procedure as the samples, but without RNA. The same volumes of RNA were analyzed by Northern hybridization using a probe for 16S rRNA (BKL54) to control for the amount of RNA in each lane. The sequencing ladder was generated using the primer BKL50, with a fragment spanning the wild type *bldD* promoter as template.


1169 +
pIJ6902
1169 +
pAU244
A C G T
0
15
30
45
60
120
0
15
30
45
60
120
P+ S1
Time post-induction (min)
Genomic *bldD* transcript
pAU244-derived *bldD* transcript
16S rRNA Northern

non–homologous tail derived from the 5' *Eco*RI site in BKL51. Both S1 nuclease assays showed the expected results. For *bldN*, transcription was seen at all time points in the control, with variations more likely due to sample loss during the S1 mapping procedure than any real change in transcription (Fig. 3.6). In contrast, *bldN* transcription in the *bldD*-induced samples clearly decreases to a minimal level by 60 minutes after *bldD* induction. In the case of *bldD*, strong transcription from the native promoter was seen at all time points in the control (Fig. 3.7). Again, this was expected; the 1169 strain has a point mutation in the coding region of *bldD*, resulting in an unstable protein, which in turn results in deregulated transcription from the *bldD* promoter. Production of active protein in the 1169 + pAU244 strain resulted in a decrease in the native transcript correlating to production of the plasmid-derived *bldD* transcript. This was also expected; BldD represses its own transcription, and induction of wild type protein from the plasmid would result in repression of the chromosomal promoter. The sequencing ladder was generated using as template a fragment that spanned the wild type *bldD* promoter. This allowed determination of the size of the *ptipA*-dependent transcript but could not be used to show that the plasmid-derived transcript originated from the promoter *ptipA*.

3.2.5.2 Hybridization of microarrays and analysis of results

Once the appropriate transcription patterns of *bldN* and *bldD* had been confirmed, the RNA samples were then used to synthesize fluorescently-labelled cDNA, which was hybridized to the microarrays. The same cDNA labelling technique was used as with the previous attempt, and once again transcription levels in the pAU244 strain were compared to those in the control (pIJ6902) strain to remove from the analysis any genes not relevant to the study (see section 3.2.3). The results show that this attempt, in a *bldD* mutant background, was much more successful than the last, in a wild type background (see section 3.2.3). Many genes were found that, as expected, were repressed as *bldD* was induced (Fig. 3.8). Surprisingly, several genes appeared to be induced by BldD, with the most significant induction seen for *sigQ* (section 3.8). Since there were many putative targets, only genes with a roughly greater than two-fold change in expression were examined, and the list was further narrowed by gene product function. Genes encoding proteins with regulatory functions, or likely to be involved in development, were given higher priority in the selection of genes for further study. Even after restricting the results to potentially interesting genes, there were too many putative targets to characterize within the scope of this thesis. Only known BldD targets and the genes chosen for biochemical analysis are listed in Tables 3.2 and 3.3, and the data from the microarray analysis illustrated in Figure 3.9. Descriptions of and microarray data from the remaining genes identified is found in Appendix C. The $\log_2$ fold change in expression of the genes of interest was plotted for each of the arrays (figures referred to within individual sections); this provides a clearer view of the degree of change. The simple fold change for a repressed gene will be between zero and one (0 to +1), while the fold change for an activated gene will be any number above one. The $\log_2$ conversion results in the same numerical distance for the same degree of change, with repression showing below zero and activation showing above zero.

Figure 3.8

Microarray analysis of global gene expression changes in response to *bldD* induction in a *bldD* mutant genetic background. The clustered image is a representation of microarray results generated by hybridization of labelled cDNA, synthesized from RNA from the *S. coelicolor bldD* mutant strain 1169, onto the *S. coelicolor* microarray. The two strains used were 1169 containing either the control parent plasmid pIJ6902 or the *bldD*-containing plasmid pAU244. Samples harvested prior to induction (to which all induced samples were hybridized) were labelled green, while samples harvested post-induction were labelled red. Repressed genes appear more green while induced genes appear more red. Grey boxes indicate spots for which no data were available. Affected genes are identified by gene name (if available) or SCO number. Search parameters limited the genes identified to those with a two-fold or greater difference from the uninduced sample, on a minimum of three arrays, although some known targets were added manually. Previously published targets are identified with an asterisk (*), and *bldD* is marked with ‡.

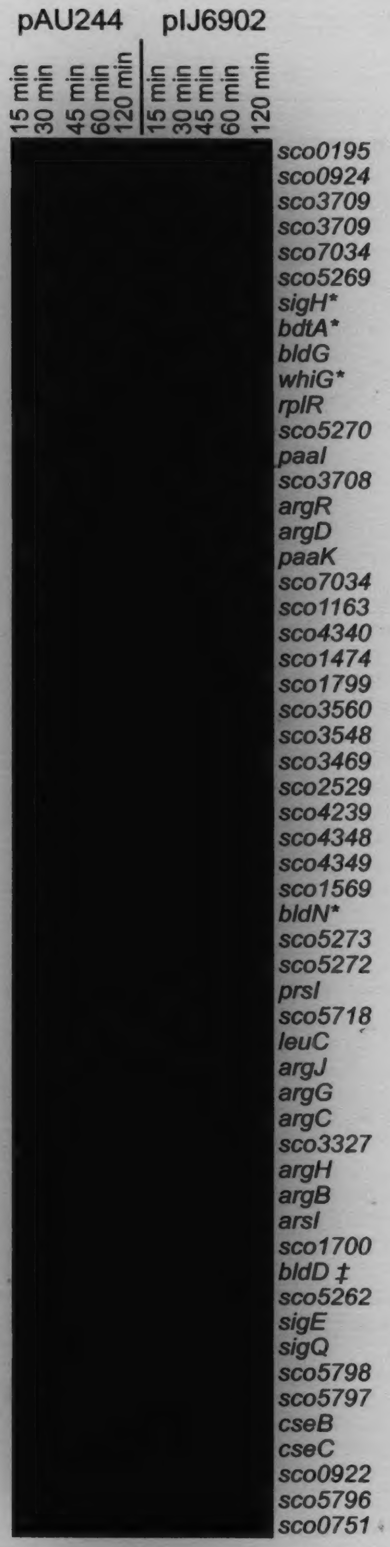

pAU244
pIJ6902
15 min
30 min
45 min
60 min
120 min
15 min
30 min
45 min
60 min
120 min
sco0195
sco0924
sco3709
sco3709
sco7034
sco5269
sigH*
bdtA*
bldG
whiG*
rplR
sco5270
paaI
sco3708
argR
argD
paaK
sco7034
sco1163
sco4340
sco1474
sco1799
sco3560
sco3548
sco3469
sco2529
sco4239
sco4348
sco4349
sco1569
bldN*
sco5273
sco5272
prsI
sco5718
leuC
argJ
argG
argC
sco3327
argH
argB
arsI
sco1700
bldD ‡
sco5262
sigE
sigQ
sco5798
sco5797
cseB
cseC
sco0922
sco5796
sco0751

Table 3.2 BldD-induced changes in gene expression of known and newly identified *bldD* target genes

Gene name	Cosmid ID	SCO number	pAU244						pIJ6902				
			t=15	t=30	t=45	t=60	t=120		t=15	t=30	t=45	t=60	t=120
bdtA	SCE68.26c	SCO3328	0.1	-0.21	-0.17	0.05	0.01		0.28	0.1	-0.04	-0.13	0.13
bldD	SC9C5.13	SCO1489	0.31	1.42		1.8	1.65		0.77	0.39	0.45	-0.01	0.48
bldN	SCE68.21	SCO3323	0.33	0.7	-0.21	-0.84	-1.33		0.91	1.52	0.29	0.39	0.31
sigH	2SC7G11.05c	SCO5243	0.04	-0.13	-1.00	-1.13	-1.98		-0.10	0.11	0.00	0.5	-0.01
whiG	SC2E1.38	SCO5621	-0.21	0.27	-0.72	-0.30			0.08	0.25	-0.47	-0.39	
**arsI*	SCE25.08c	SCO3067	-0.19	-0.07	0.01		0.25		0.91	0.32	0.89	-0.1	0.61
**prsI*	SCE25.07c	SCO3066	-0.13	-0.51	-0.73	-0.92	-0.73		0.23	0.02	0.01	-0.02	0.33
**sigQ*	2SCK8.34	SCO4908	2.58	2.62	1.05	4.38	2.72		0.01	0.45	0.21	0.21	-0.11

Empty cells indicate unavailable data.

Numbers represent the $\log_2$ fold change in expression of genes in the indicated samples relative to the uninduced sample. Positive numbers indicate induction while negative numbers indicate repression of transcription.

* putative BldD target identified by the microarray analysis

Table 3.3 Description of known and newly identified BldD target genes

Gene name	**Cosmid ID**	**SCO number**	**Gene product function (known or proposed) ****	**Importance in this study**	**BldD binding sequence Consensus match**
bdtA	SCE68.26c	SCO3328	Putative DNA-binding protein	BldD target (repressed) (Elliot *et al*, 2001)	AGTGA GCACG (13) TCACC
bldD	SC9C5.13	SCO1489	DNA binding protein	Regulator of morphological differentiation and antibiotic production (Merrick, 1976)	AGTAA (12) TCACA
bldN	SCE68.21	SCO3323	ECF sigma factor	Regulator of morphological differentiation (Bibb *et al*., 2000), direct BldD target (repressed) (Elliot *et al*, 2001)	AGTGC (2) GCACG, GGTGA (13) GCACG
sigH	2SC7G11.05c	SCO5243	Sigma factor	BldD target (repressed, spatially regulated) (Kelemen *et al*, 2001)	AGTGA (2) GCACG
whiG	SC2E1.38	SCO5621	Sigma factor	Regulator of sporulation (Chater, 1975), BldD target (Elliot *et al*, 2001)	AGTGA TCACC, AGTCA (8) TCACG
**arsI*	SCE25.08c	SCO3067	Putative anti-anti-sigma factor	BldD target (repressed)	AGTGC (12) TCACG
**prsI*	SCE25.07c	SCO3066	Putative anti-sigma factor	BldD target (repressed)	GGTCG (8) ACACC, GGTGG (3) GGTGAGGTGA (11) GCACC
**sigQ*	2SCK8.34	SCO4908	ECF sigma factor	BldD target (activated)	See Figure 3.39

* putative BldD target identified by the microarray analysis. Note: *arsI* was not identified by the microarray analysis, but was included because it is transcribed in an operon with *prsI*.

** ECF = extracytoplasmic function

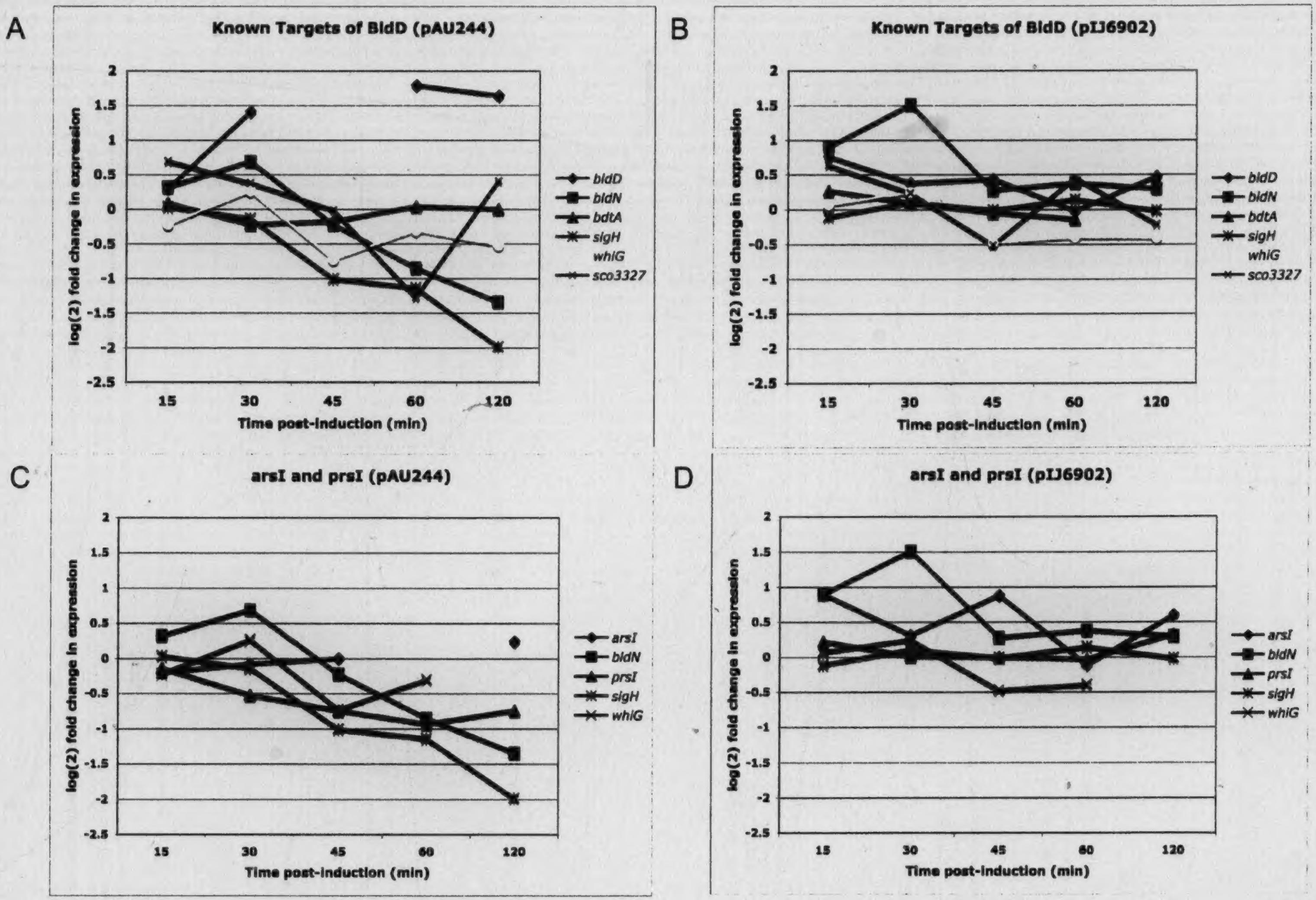
Figure 3.9

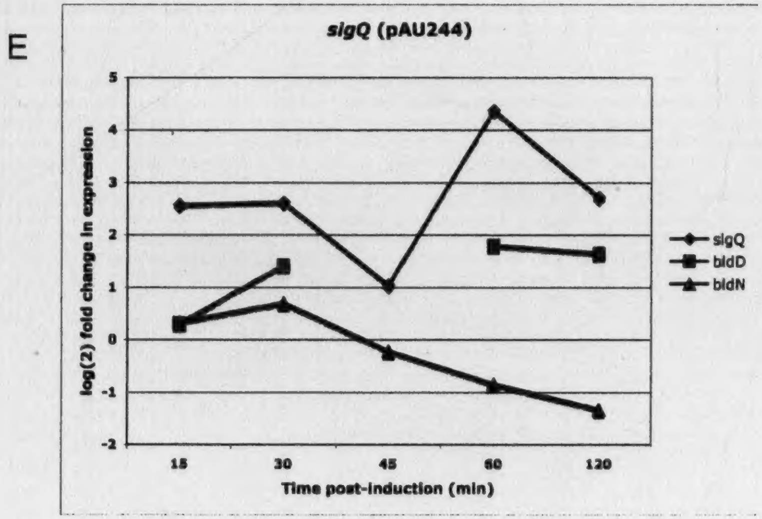
Microarray results (see Table 3.3), plotted as the $\log_2$-fold change in expression relative to the uninduced sample at each time point post-induction. Values above zero indicate induction and values below zero indicate repression. The plots on the left side of each page (A, C and E) represent the samples harvested from the *bldD* mutant strain containing inducible *bldD* (1169 with pAU244), while the plots on the right side of each page (B, D and F) represent the samples harvested from the *bldD* mutant strain containing the control plasmid (1169 with pIJ6902). Data were not available for all time points for all genes; gaps in the data are shown by breaks in the lines.

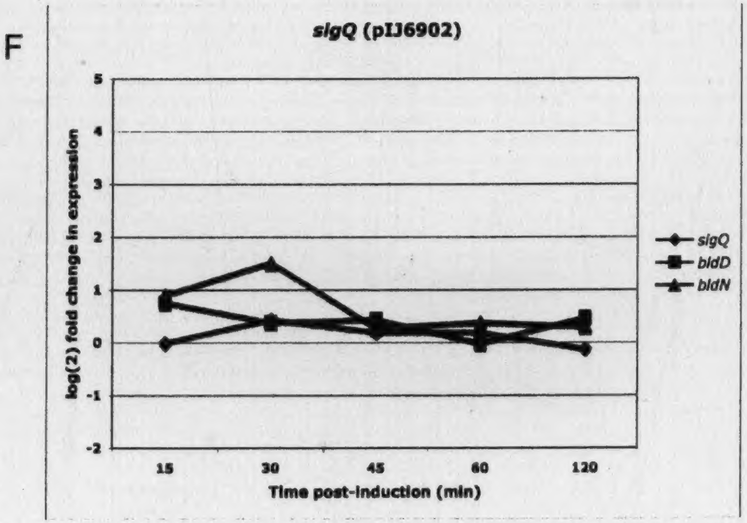
A) and B) Published BldD targets. *sco3327* is transcribed as an operon with the direct BldD target *bdtA* (Elliot, 2000).

C) and D) Newly identified BldD targets *prsI* and *arsI*, along with known targets *bldN*, *sigH* and *whiG*, for comparison.

E) and F) Induced target *sigQ*, with the repressed target *bldN* included for comparison. *bldD* transcription is induced as part of the experiment.


A
Known Targets of BldD (pAU244)
log(2) fold change in expression
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
15
30
45
60
120
Time post-induction (min)
bldD
bldN
bdtA
sigH
whiG
sco3327
B
Known Targets of BldD (pIJ6902)
log(2) fold change in expression
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
15
30
45
60
120
Time post-induction (min)
bldD
bldN
bdtA
sigH
whiG
sco3327
C
arsI and prsI (pAU244)
log(2) fold change in expression
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
15
30
45
60
120
Time post-induction (min)
arsI
bldN
prsI
sigH
whiG
D
arsI and prsI (pIJ6902)
log(2) fold change in expression
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
15
30
45
60
120
Time post-induction (min)
arsI
bldN
prsI
sigH
whiG


E
sigQ (pAU244)
log(2) fold change in expression
5
4
3
2
1
0
-1
-2
15
30
45
60
120
Time post-induction (min)
sigQ
bldD
bldN


F
sigQ (pIJ6902)
log(2) fold change in expression
5
4
3
2
1
0
-1
-2
15
30
45
60
120
Time post-induction (min)
sigQ
bldD
bldN

Transcription patterns of all known *bldD* targets were examined first, as a final check to validate the experiment (Fig. 3.9, A and B). Transcription of *bldD* clearly increased, stabilizing by 60 minutes post-induction. Transcription of *bldN* and *sigH* decreased, as expected, demonstrating that known targets can be identified through this analysis. Transcription of *bldN* as measured by microarray analysis correlated well with the transcription pattern seen by S1 nuclease mapping (Fig. 3.6). However, transcription of *bdtA* and *whiG* remained constant after *bldD* induction, and *sco3327* (co-transcribed with *bdtA*) showed first repression, then induction. These results both confirm the necessity to verify the microarray results and indicate that not all of the targets of BldD regulation can be identified by a single microarray experiment.

Since opportunities to repeat the microarray experiment were limited, a decision was made to follow up on the most convincing of the genes identified. Those chosen included *prsI* and *sigQ* and the follow-up involved biochemical analysis. For several other potential targets (see Appendix C), follow-up was limited to examination of the sequence upstream of the coding region. All sequence and annotation information was found on ScoDB (Bentley *et al.*, 2002). Results are discussed in subsequent sections.

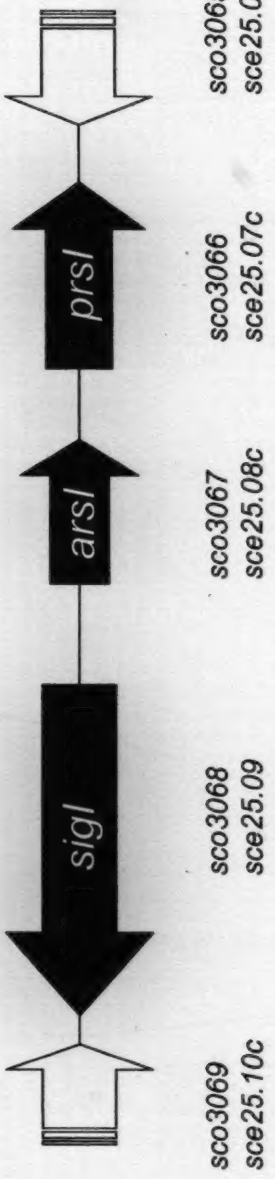
3.3 *prsI* is a target of BldD regulation

3.3.1 Sequence analysis of the *prsI* promoter region

prsI (*sce25.07c*, *sco3066*), encoding a putative anti-sigma factor (Fig. 3.10, Table 3.3) was one of the targets identified by the initial round of analysis

Figure 3.10

Gene structure in the chromosomal area surrounding *prsI*. The genes on the far left and right (*sco3069* and *sco3065*) are included for reference; *sigI*, *arsI* and *prsI* are discussed in this work. SCO numbers and cosmid designations are included beneath the relevant genes; names are included where possible. With the exception of *sco3069* and *sco3065*, for which only partial genes are shown, gene sizes and the intergenic regions are drawn to scale.


sco3069
sce25.10c
sigI
sco3068
sce25.09
arsI
sco3067
sce25.08c
prsI
sco3066
sce25.07c
sco3065
sce25.06

from the microarray experiment in which *bldD* was inducibly overexpressed in the *bldD* mutant, *S. coelicolor* 1169. The predicted product of the *prsI* gene has homology to the kinase domain of RsbW of *Bacillus subtilis*, an anti-sigma factor that regulates the activity of the primary stress-response sigma factor of *B. subtilis*, σ^B (Hughes and Mathee, 1998) as determined using the NCBI Conserved Domain Database (http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml). As shown in Figure 3.9 C and D, no significant change was seen in *prsI* transcript levels as the control strain was induced, but a clear reduction in *prsI* transcript levels was seen as *bldD* was induced in the test strain (data in Table 3.2). As a first step in verifying that *prsI* is a direct target of BldD, the sequence upstream of the *prsI* start codon was checked for the presence of the BldD binding consensus AGTgA (n_m) TCACc (Elliot *et al.*, 2001). Initial examination found only one consensus match - GGTGG (3 bp) GGTGA GGTGA (11 bp) GCACC (Fig. 3.11). The *prsI* start codon immediately precedes the second half of the consensus. Later examination, after initial transcript mapping, found a second BldD binding consensus match further upstream - GGTCG (8 bp) ACACC.

3.3.2 Analysis of *prsI* transcription patterns

The position of BldD binding to all known targets was consistent with the expected location of repressor binding. In order to see if this was the case for *prsI* as well, the transcription start point(s) had to be determined. Both S1 mapping and primer extension were used to determine the transcription start point(s) of *prsI* and to compare the pattern of transcription in wild type and *bldD*

Figure 3.11

Sequence of the *prsI* coding region and surrounding area, including part of the *arsI* coding region. Only the coding strand is shown. Start and stop codons are marked in bold italics, coding regions (*prsI* and part of *arsI*) are shown in bold, and the *prsI* transcription start points (*prsIp1* and *prsIp2*) are marked in italics with a bent arrow above. The sequences showing similarity to the BldD binding consensus sequence are marked in red, with a red arrow above indicating the first (forward arrow) or second (reverse arrow) portion of the half site. The locations of oligonucleotide primers are marked with an arrow indicating the strand they are derived from; forward primers are marked above the sequence, and reverse primers are indicated below. Wavy lines at the end of some primers indicate engineered restriction sites; the colour of the restriction enzyme is matched to the colour of the primer. Where primers include sequence from two lines in the figure, the arrow and primer name are colour coded. As well, in crowded areas, some primers are colour coded to minimize confusion with other labelled features.

*Bam*HI CGA20

ggaagccgcg gggggcggag tgcatttggt ggccatgcag ccggtagtgg

cgcgcgtgtt cgagatcacg ggggccgaag cggtcttcac ccttcatgac

acgcttgcgg ccgccctggc cgacgcgtcc gac_tga_gccg tccttgtccg

CGA44

gccgacggag gggtgcgtgt gcttcccgga gcgggtgttt tcgccccgcg

CGA49

ccgggtttcg tgcgcaatac aggggtgttc gtgcgatccg **gtcgggcagg**

prslp2 CGA43

agacaccct*g* aaccggctgt gtacgacatg gctgtctacg actttgaaac

prslp1 CGA51 CGA50 *Nde*I CGA64

*g*tgaactggt ggatcggtga ggtgaagcgc tg**_atga_gcac cacccggccc**

CGA18/CFU1 *Bgl*II

tacccgccgg gcgaccgcgg tccggagccc ggcggcgctt ccgggacggc

gtccgcgccc gcctccgcgt ccggggaggg cggcggacgg caggcccgca

CGA42

ggctgagctt cgaggacgcg agcggggtcg tcccgctggc ccgtgacttc

CGA67

acccgcgagg cgctgtacgc ctggggctgg ctgccctccg ccaccgcgga

ccagcgcgcc gccgcggagg acgtcctgct cgtggtctcc gagctggtca

ccaacgcgtg tctgcacgcg gagggcccgg acgagctgtc gatcacctgt

gagaagaagg tgatccggct ggaggtctcc gaccgcggta cgggtcagcc

ggcgccgcgg accccgcacc gcgcggggcg ccccggcgga cacggcatgt

tcatcgtgca gcggctctgc ctggactggg gcgtcgtacg gacgccgggg

gtcgcgggca agcgggtgtg ggcggagctg ggcgccccgg cg_tag_gcctg

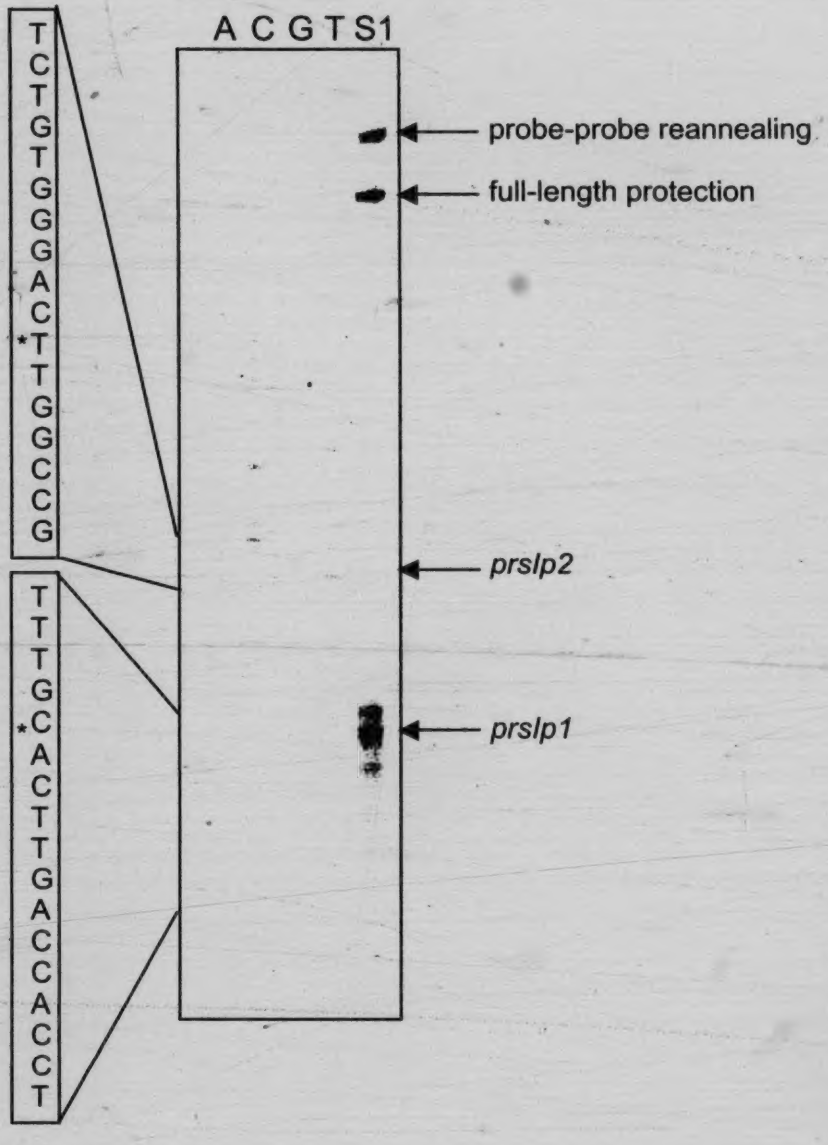
cggcgctcgg gcgggcaccc gcccgtccgc ccacccgtct cggtcggcca

CGA65

gcaggccgcc gtcccggggg ccgcgccccg gcccgcagag cggaccgcgc mutant strains over different time points in development. Cultures for the RNA time courses were grown on rich solid medium (R2YE). RNA was isolated from wild type (916) and *bldD* mutant (1169) *S. coelicolor* strains at seven time points, ranging from very early vegetative growth (prior to formation of aerial mycelia) to a time when the wild type strain had reached late sporulation. Integrity of the RNA and absence of any DNA contamination was confirmed by agarose gel electrophoresis, and quantification of the RNA samples was confirmed by Northern analysis using a probe for 16S rRNA (BKL54). A number of different probes were designed for the S1 mapping; the probe which allowed clearest visualization of protected fragments was CGA20-CGA67 (Fig. 3.11, Table 2.4), a 439 bp probe that extended from -301 to +138 relative to the *prsI* translation start codon and which contained a ten base pair non-homologous tail to distinguish between probe:probe reannealing (would result in a 439 bp band) and full-length protection (would result in a 429 bp band). In initial attempts, three protected bands were observed, one appearing to correspond to a full-length protected probe fragment and two others corresponding to potential transcription start points (not shown). Resolution of the full-length protection band was difficult because of the length of the probe, so the purified PCR product was cloned into pCR®2.1 TOPO (creating pAU248), then the S1 probe amplified using the M13 reverse primer (Invitrogen) (within the vector) and CGA67. This resulted in the formation of a 539 bp probe with a 100 bp non-homologous tail and allowed easy distinction between probe-probe reannealing and full-length protection (Fig. 3.12). Use of this probe resulted in visualization of two shorter than full-length

Figure 3.12

High resolution S1 nuclease mapping of the *prsI* transcription start points. The 539 bp probe was amplified from pAU248 using the primers M13 reverse (primer binding site located inside the vector sequence) and CGA67 (binds within the *prsI* coding region), then end-labelled. The bands corresponding to two transcription start points, full length protection of the probe resulting from a read-through transcript, and probe-probe reannealing are indicated by arrows. A sequencing ladder generated using the primer CGA67 was used to determine the location of the transcription start points. A portion of the surrounding sequence is found to the left of the ladder, with the most likely transcription start points marked with asterisks (*).

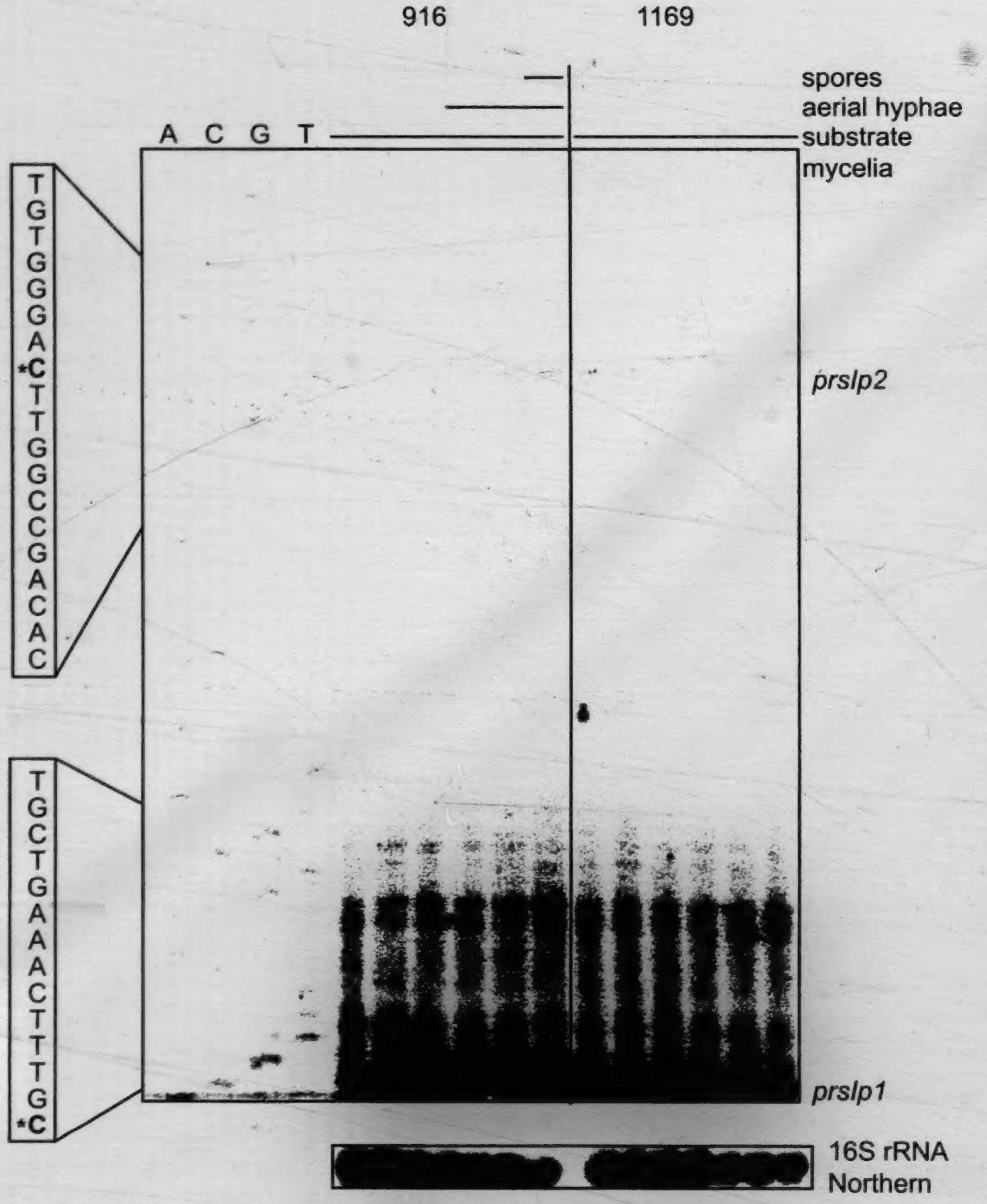

A C G T S1
probe-probe reannealing
full-length protection
prslp2
prslp1
T
C
T
G
T
G
G
G
A
C
*T
T
G
G
C
C
G
T
T
T
G
*C
A
C
T
T
G
A
C
C
A
C
C
T

protected bands suggesting the existence of two transcription start points, one 32 bp (*prslp1*) and one 73 bp (*prslp2*) upstream of the *prsl* translation start codon. Since other additional, fainter bands were also seen, the results were confirmed using primer extension experiments.

A number of different primers were used for primer extension. The probe that was used to successfully map the transcription start points was CGA18, a primer that overlaps the start codon of *prsl*. Extension from this probe resulted in three bands, two of which correlated with the previously mapped *prslp1* and *prslp2* (Fig. 3.13) and one larger band which was presumed to be falloff from the read-through transcript from the upstream *arsl*. This larger band will be discussed further in section 3.4.2. As was the case for previously studied *bldD* targets, BldD repression of both *prslp1* and *prslp2* was indicated by elevated levels of transcript at the earliest time point in the *bldD* mutant compared to the wild type strain, indicating developmental regulation of *prsl* transcription. With respect to the two *prsl*-specific promoters, signal from *prslp2* was absent in the very first time point tested in the wild type, and transcription from *prslp1* in the wild type was less than that from the earliest vegetative time point in the *bldD* mutant. As stated above, identification of *prslp2* and evidence of *bldD* dependence prompted closer analysis of the DNA sequence surrounding the *prslp2* promoter region. A potential match to the BldD binding consensus sequence was found and led to further investigation of this site through DNA-protein interaction studies (section 3.3.3). Using primer extension, the shorter product corresponding to the transcript originating from *prslp1* was much more

Figure 3.13

Primer extension to analyze levels of expression from the *prslp1* and *prslp2* promoters in RNA isolated from wild type (916) and *bldD* mutant (1169) *S. coelicolor* strains over a time course of development. The primer used was CGA18, which overlaps the *prsl* start codon. The stage of development is indicated by the horizontal lines above the bands. Transcription start points are indicated in the sequence by bold type and an asterisk (*), and on the right by a text label. The same volumes of RNA were analyzed by Northern hybridization using a probe for 16S rRNA (BKL54) to control for the amount of RNA in each lane. The apparent doublet in *prslp1* is the result of an abrupt change of acrylamide concentration at the bottom of the sequencing gel, and not the result of two separate transcripts. The lower band likely represents labelled probe. The probe used (CGA18) overlaps the start codon of *prsl*, and *prslp1* is only 32 bp upstream of the translational start of *prsl*.


916
1169
spores
aerial hyphae
substrate
mycelia
A C G T
T G T G G G A *C T T G G C C G A C A C
prslp2
T G C T G A A A C T T T G *C
prslp1
16S rRNA
Northern

difficult to visualize, being only 28 bp longer than the primer itself. This transcription start point was found immediately upstream of the complex putative BldD binding site, that is the one with three potential upstream half sites and only a single downstream half site, and 32 bp upstream of the *prsI* start codon. In both the S1 protection and primer extension analyses, transcription levels from *prsIp2* were much lower than those from either *prsIp1* or the read-through transcript (Figs. 3.12 and 3.13). *prsI* transcription patterns were confirmed on two independent time courses.

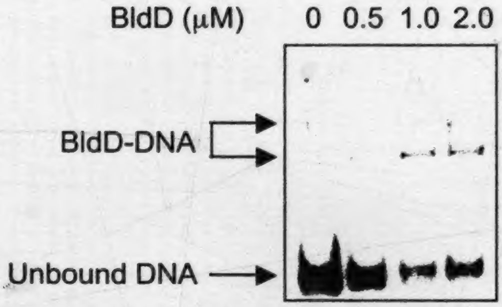
3.3.3 Binding of BldD to the *prsI* promoter region

Prior to testing the ability of BldD to shift the *prsI* promoter, an electrophoretic mobility shift assay (EMSA) was performed to test BldD binding to the *bldD* promoter using the method developed by M. Elliot (Elliot and Leskiw, 1999). This was done as a positive control of the methodology to be used for the testing of BldD binding to the newly identified targets (Fig. 3.14). His$_6$-BldD (Elliot and Leskiw, 1999) purified by M. Elliot in 1999 was used in all assays requiring BldD. The *bldD* probe used was MAE16-4 (77 bp). Binding of BldD to its own promoter was seen and the K_D was calculated to be 0.4 μM, comparable to what has been found previously (Fig. 3.14) (Elliot and Leskiw, 1999). The experiment demonstrated BldD binding to the *bldD* promoter and served both as a trial of technique and to show that the protein was still active.

Comparison of transcript levels in the wild type versus *bldD* mutant strain suggested that BldD regulates the transcription of *prsI*. In the absence of additional information, this alone does not distinguish between direct and indirect

Figure 3.14

EMSA demonstrating BldD binding to its own promoter. The probe used was MAE16-4 (77 bp), which had been cloned into pUC119 then purified by digestion with *Eco*RI and *Bam*HI and gel purification. The probe was end-labelled, then redissolved to 1 ng/µL. The probe was then incubated with increasing amounts of BldD and loaded on a running 8% mini-PAG with glycerol to separate the bound DNA from the unbound. Bound and unbound DNA are indicated by arrows, and the amount of BldD added is indicated across the top.


BldD (μM)
0
0.5
1.0
2.0
BldD-DNA
Unbound DNA

regulation of the target genes. Therefore, EMSAs were carried out to confirm that BldD binds upstream of *prsI*. The intergenic region between *arsI* and *prsI*, along with the probes used for EMSAs, is detailed in Fig. 3.15. All EMSA reactions were carried out using the previously purified His_6-BldD (Elliot and Leskiw, 1999), incubated with 1 ng of labelled probe in the presence of the non-specific competitor poly d(I-C), then loaded on a running glycerol-containing 8% polyacrylamide gel. Separate probes were initially used for testing binding to the upstream and downstream putative binding consensus sequences. The first probe, used for the downstream binding site (downstream of *prsIp1*), was CGA51-18, a 41 bp PCR product extending from 26 bp upstream of the ATG to 15 bp inside the coding region, giving 1 bp upstream and 6 bp downstream flanking the BldD binding site. Use of CGA51-18 to demonstrate binding to the putative BldD binding site downstream of *prsIp1* - GGTGG (3 bp) GGTGAGGTGA (11 bp) GCACC - showed very weak binding, such that none of the probe was visibly bound (Fig. 3.16, A). However, BldD bound strongly to a longer probe, CGA43-42, that extends more than 50 bp beyond the consensus in either direction (Fig. 3.16, B); the K_D of BldD binding was calculated to be 1.0 μM (see section 2.6.1). BldD binding was specific, as demonstrated by addition of a 500-fold excess of unlabelled probe; although the non-specific interaction is not entirely eliminated, the BldD-DNA is reduced to a barely perceptible level. This indicates that BldD needs a minimum amount of flanking sequence in order to bind. The flanking sequence could also be required to allow a change in DNA conformation either to allow binding or immediately after binding. The smallest

Figure 3.15

Strategy for EMSA of BldD binding to the *prsI* promoter region.

A) Schematic of the *arsI/prsI* intergenic region with primers and potential BldD binding sites marked. *prsIp1* and *prsIp2* are indicated on both parts of the diagram (bent arrows). Determination of the transcriptional start sites for *sigI* and *arsI*, designated p_{sigI} and p_{arsI}, is discussed later in sections 3.5.2 and 3.4.2, respectively. The expanded region, drawn to scale, contains the potential BldD binding sites and primers used for the EMSA experiments. Direction of shading and arrows indicate the orientation of the BldD binding half sites. Forward primers are drawn above the line of the expanded section, and reverse primers are drawn below the line.

B) Sequence of the expanded region. Potential BldD binding sites are indicated in bold, with an arrow to show the orientation of the half-site. Transcription start points are indicated with bent arrows, and the translation start codon of *prsI* is surrounded by a box. A 29 bp gap in the sequence shown is indicated by '...$(N)_{29}$...'.

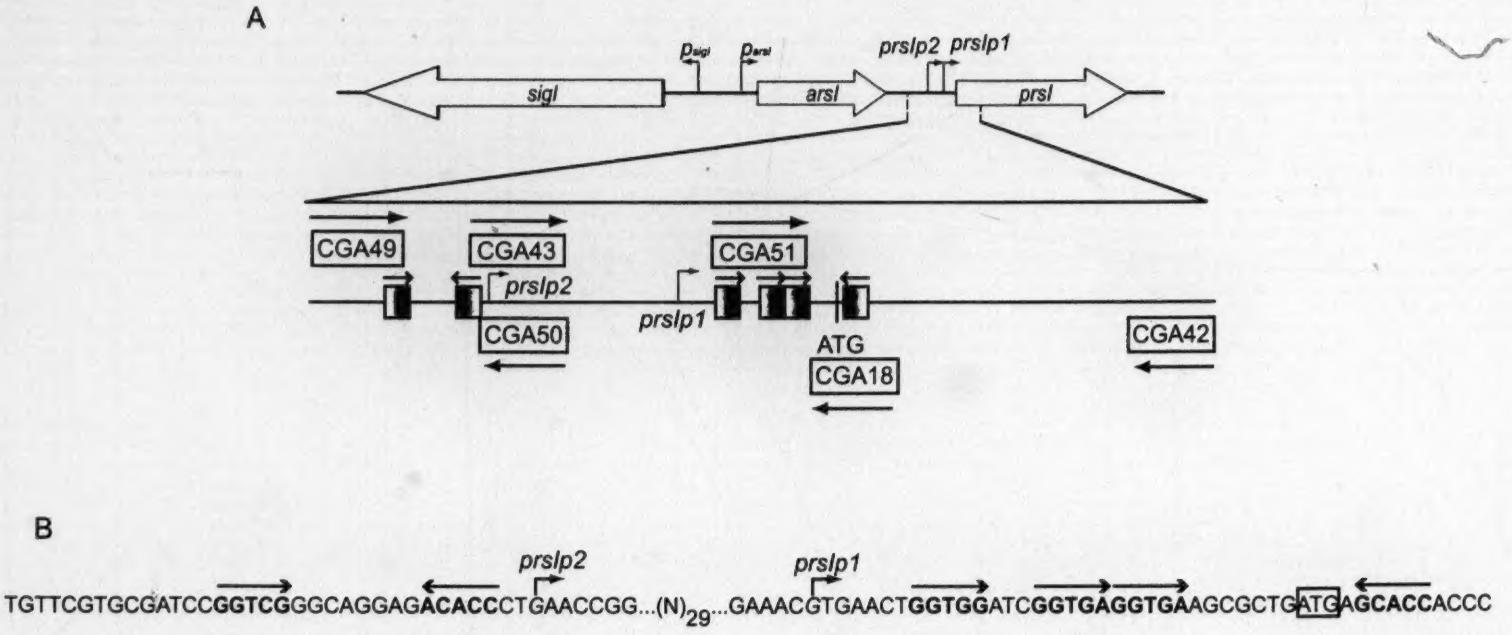
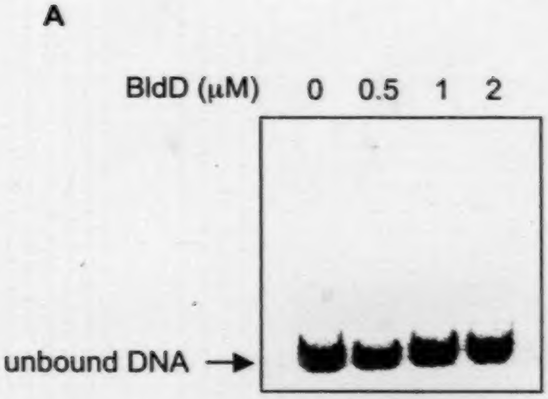

A
sigI
arsI
prsI
p_{sigI}
p_{arsI}
prsIp2
prsIp1
CGA49
CGA43
CGA51
prsIp2
prsIp1
CGA50
ATG
CGA18
CGA42
B
prsIp2
prsIp1
TGTTCGTGCGATCCGGTCGGGCAGGAGACACCCTGAACCGG...(N)$_{29}$...GAAACGTGAACTGGTGGATCGGTGAGGTGAAGCGCTGATGAGCACCACCC

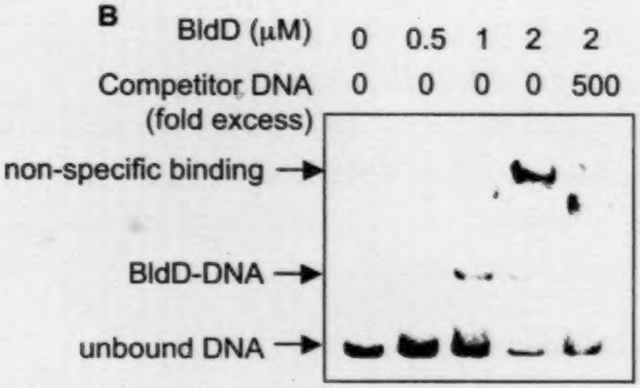
Figure 3.16

EMSA of BldD binding to the *prsI* promoter region near *prsIp1*. One nanogram of end-labelled probe was incubated with increasing amounts of purified His$_6$-BldD and loaded on a running glycerol-containing 8% polyacrylamide gel. Bound and unbound DNA are indicated by arrows.

A) Binding of BldD to the 41 bp probe CGA51-18, which has 6 bp upstream and 1 bp downstream of sequence flanking the putative BldD binding site.

B) Binding of BldD to the 162 bp probe CGA43-42, which has a minimum 52 bp flanking sequence on either side of the BldD binding site. Competitor DNA is a five hundred-fold excess of unlabelled CGA43-42.


A
BldD (μM)
0
0.5
1
2
unbound DNA


B
BldD (μM)
0
0.5
1
2
2
Competitor DNA (fold excess)
0
0
0
0
500
non-specific binding
BldD-DNA
unbound DNA

reported flanking region seen that fully supports BldD binding had a minimum of 10 bp flanking region (Kim *et al.*, 2006), suggesting a parameter for future probe design. The probe designed to test BldD binding to the putative binding site upstream of *prslp2*, GGTCG (8 bp) ACACC, was CGA49-50, a 51 bp PCR product. Binding to this probe was weak, but addition of a 500-fold excess of unlabelled probe eliminated the BldD-DNA band indicating that what binding occurred was due to specific interactions and not to non-specific binding (Fig. 3.17, A). The binding might have been stronger with a longer probe, though there was a minimum of 15 bp on either side of the consensus.

Recent work has shown that BldD binds very strongly to only one of the binding sites in the promoters of *bldN* and *whiG* if the two sites in each promoter are amplified and tested separately (Kim *et al.*, 2006). This was also seen in *prsI*, to the extent that one site was bound very strongly and the other very weakly. To see if there was any change in the strength of binding of BldD to both sites as compared to the downstream site, a longer probe was designed. CGA49-42 encompasses both of the putative BldD binding sites and extends from 108 bp upstream of the *prsI* start codon to 88 bp inside the coding region. As with the other EMSAs, 1 ng labelled probe was incubated with increasing amounts of His$_6$-BldD in the presence of poly d(I-C), then bound probe was separated from unbound probe by loading on a running, glycerol-containing 8% polyacrylamide gel (Fig. 3.17, B). Binding was seen with as little as 0.5 μM BldD, and the K_D was calculated to be 1.3 μM. Binding was not entirely eliminated with the addition of a 500-fold excess of unlabelled probe, but it was greatly reduced.

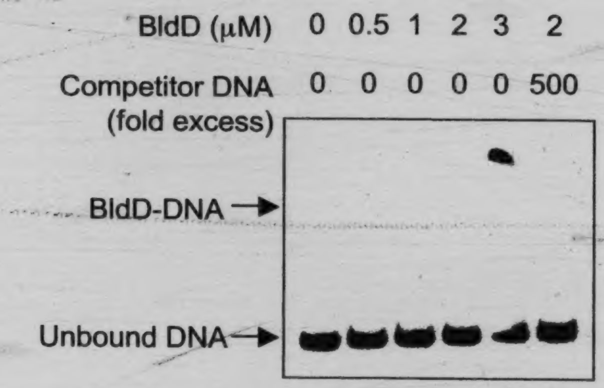
Figure 3.17

EMSA of BldD binding to the *prslp2* region of the *prsl* promoter region. One nanogram of end-labelled probe was incubated with increasing amounts of purified His_6-BldD and loaded on a running glycerol-containing 8% polyacrylamide gel. Bound and unbound DNA are indicated by arrows. Competitor DNA (in the final lane) was a 500-fold excess of unlabelled probe: the CGA49-50 and CGA49-42 DNA fragments were used as probes for A and B, respectively.

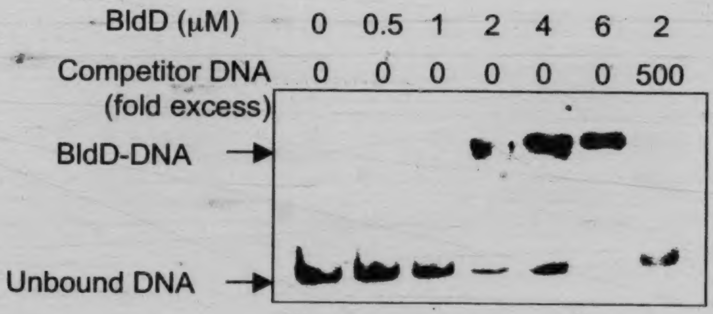
A) The 51 bp probe used (CGA49-50) spanned the BldD binding site near *prslp2*.

B) The 196 bp probe (CGA49-42) encompassed both of the BldD binding sites near *prslp1* and *prslp2* upstream of the *prsl* coding region.

A


BldD (μM) 0 0.5 1 2 3 2
Competitor DNA (fold excess) 0 0 0 0 0 500
BldD-DNA
Unbound DNA

B


BldD (μM) 0 0.5 1 2 4 6 2
Competitor DNA (fold excess) 0 0 0 0 0 0 500
BldD-DNA
Unbound DNA

This indicates that binding was due to specific interactions between BldD and its target sites. It was rather odd that the amount of unshifted probe increased at 4 μM BldD, but the experiment was not repeated so this may be an artifact of this particular gel, such as an unevenness in loading of the binding reactions.

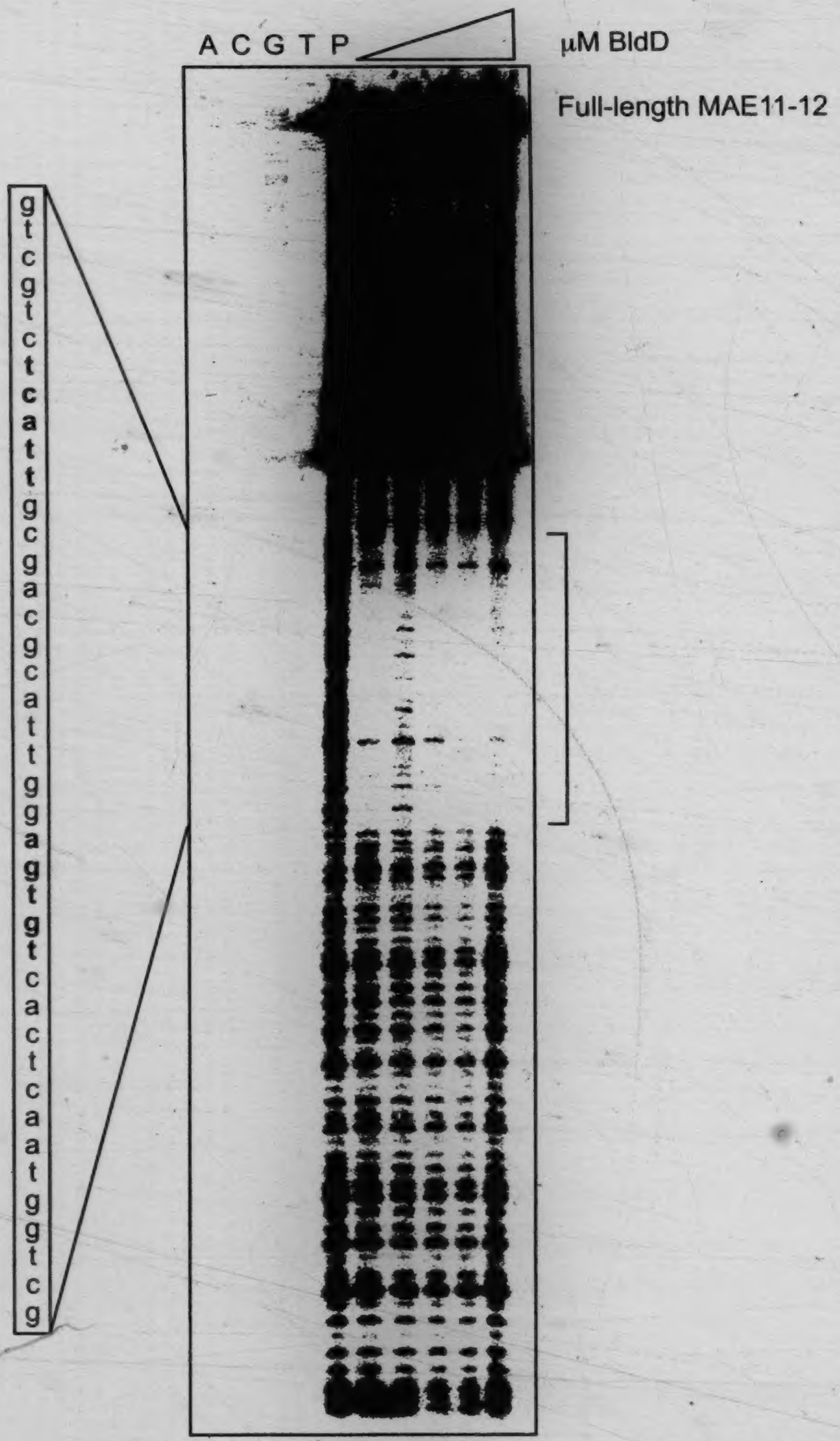
Although the EMSA with the longest probe (CGA49-42) was only performed once, all other EMSA experiments were performed at least twice. Together, they all show that *prsI* is a direct target of BldD regulation.

3.3.4 DNase I footprinting of the *prsI* promoter region

Purified His_6-BldD (Elliot and Leskiw, 1999) was used to footprint the *bldD* promoter, again to validate BldD binding to a known target. The binding reactions for DNase I footprinting were the same as those used for EMSA: incubation of labelled probe with increasing amounts of His_6-BldD (Elliot and Leskiw, 1999) in the presence of poly d(I/C). An important difference was that for an EMSA, the purified probe would simply be end-labelled and precipitated, while for a footprinting reaction, one primer only was labelled. The labelled primer was used for PCR, so that only one strand of the resulting probe was labelled. The probe used was the same as previously described - MAE11-12, a 257 bp probe which extends from 168 bp upstream of the *bldD* transcription start point to 28 bp inside the *bldD* coding region. MAE12 was end-labelled before PCR amplification and purification of the probe. The footprinting reaction was carried out exactly as described (Elliot and Leskiw, 1999; Elliot, 2000). The resulting footprint (Fig. 3.18) covered the same region as previously reported, from -29 to +13 (Elliot, 2000). Since the footprinting experiment worked with the positive

Figure 3.18

DNase I footprinting of the *bldD* promoter using the 257 bp MAE11-12 probe. Increasing amounts of His$_6$-BldD (0.5, 1, 2, 3 and 4 μM) were incubated at 30°C for 20 minutes with 0.06 pmol probe, labelled on the template strand (using end-labelled oligonucleotide MAE12), then 0.05 U of DNase I was added and allowed to react for 10 seconds. Digestion products were resolved on a 6% sequencing gel, alongside a sequencing ladder (labelled A C G T) generated with unlabelled MAE12 oligonucleotide as primer. P indicates probe incubated without BldD, then treated with DNase I. The published footprint (Elliot and Leskiw, 1999; Elliot, 2000) is shown with a bracket and in the sequence beside the gel. The sequence matching the BldD consensus binding sequence is indicated in bold type.

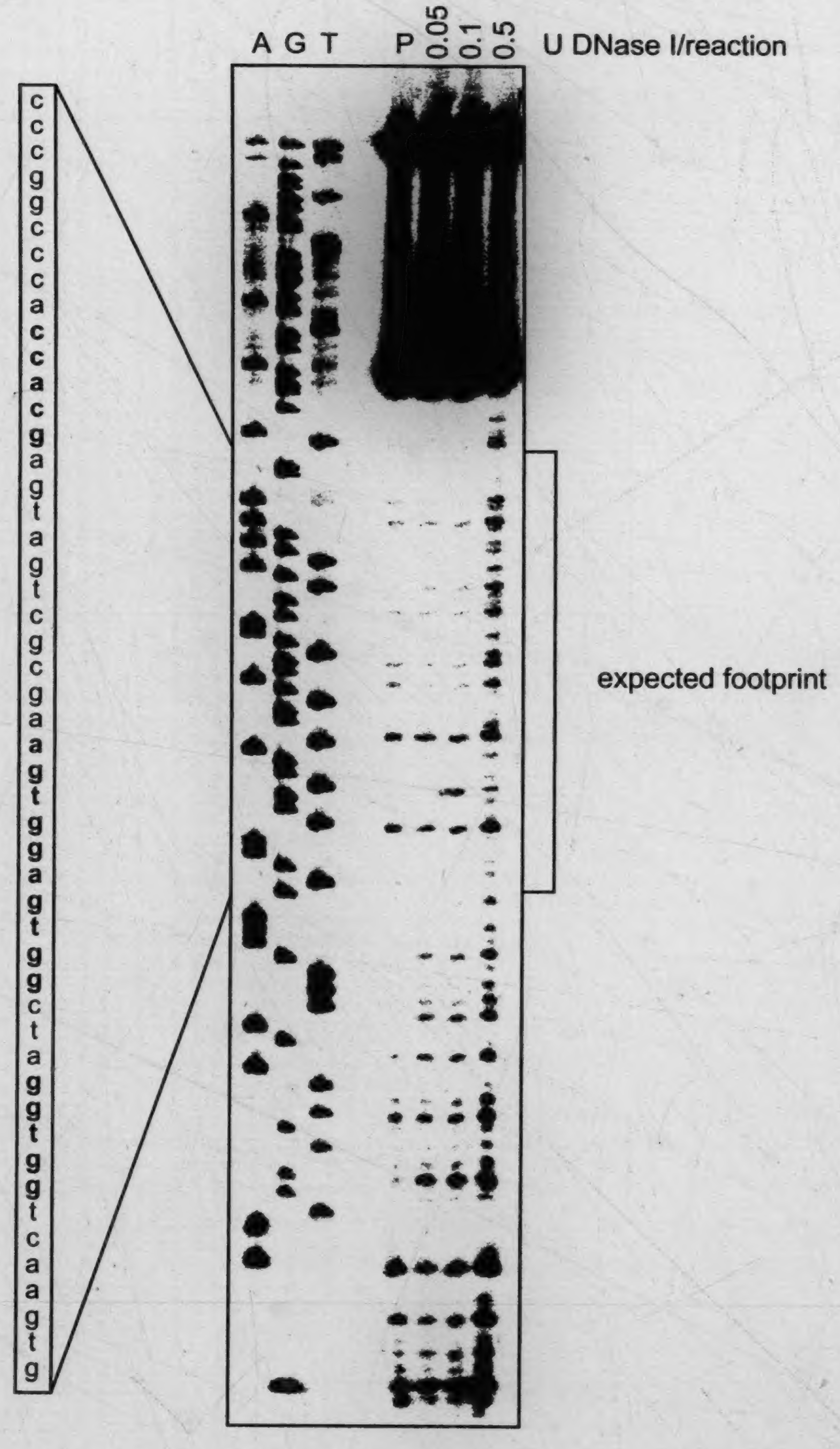

A C G T P
µM BldD
Full-length MAE11-12
gtcgtc**tcatt**gcgacgcattgg**agtgt**cactcaatggtcg

control, the same technique was then used to footprint two of the newly discovered BldD targets.

Attempts were first made to footprint BldD binding to the *prsI* promoter. The probe used was CGA43-42, a 162 bp PCR product which extends from -4 relative to *prsIp2* to 87 bp inside the *prsI* coding region (Figs. 3.11, 3.15), including the BldD binding consensus match near *prsIp1* but not the weaker match upstream of *prsIp2*. Unfortunately, the region was never successfully footprinted due to insurmountable problems with probe degradation. Regardless of which strand was labelled (primer labelled prior to PCR), a large amount of degradation of the probe was seen (Fig. 3.19). The location of the degradation was consistent, 29-30 bp inside the coding region on the coding strand, and 31-32 bp upstream of the start codon (near *prsIp1*) on the non-coding strand. A range of DNase I concentrations was used in one experiment, but this had no effect whatsoever on the degradation product (Fig. 3.19). Many attempts were made to fix or circumvent the problem, which was also seen when CGA21-19 was used to footprint BldD binding the promoter region of *arsI*. A more detailed description is found in section 3.4.4. Due to the nature of the sequence surrounding the BldD binding consensus sequences in the *prsI* promoter, it was impossible to design a shorter probe for footprinting, such as the probe successfully used for footprinting of BldD binding to the *arsI* promoter (see section 3.4.4). Given that the *prsI* probe is shorter (162 bp vs. 257 bp, Fig. 3.18) than the *bldD* probe, which was used successfully despite a similar level of degradation, probe length is clearly not the whole story. The batch of protein

Figure 3.19

Attempt at DNase I footprinting of BldD binding to *prslp1* using the 142 bp probe CGA43-42. The probe was labelled on the coding strand. Increasing amounts of DNase I were used to treat 0.06 pmol of probe that had been incubated with 3μM His$_6$-BldD. P indicates probe incubated without BldD, then treated with DNase I. Digestion products were resolved on a 6% sequencing gel, alongside a sequencing ladder generated with CGA43. The expected footprint (based on sequence analysis) is shown with a bracket. The BldD binding sequence is typed in bold within the sequence to the left of the figure.


A G T
P
0.05
0.1
0.5
U DNase I/reaction
cccggcccaccacgagtagtcgcgaagtggagtggctaggtggtcaagtg
expected footprint

was also not an issue, as the first trial with the *prsI* probe was done in the same time frame as the successful *bldD* footprinting. The EMSA results shown above suggest the strength of BldD binding to the target DNA may provide some explanation, as BldD does not bind the *prsI* promoters as tightly as the *bldD* promoter.

3.4 *arsI* is a target of BldD regulation

3.4.1 Sequence analysis of the *arsI* promoter region

Examination of the gene structure in the area of *prsI*, courtesy of SCO-DB, revealed a second gene (*arsI*, *sce25.08c*, *sco3067*) which, because of its position directly upstream of *prsI* and because of the presence of a read-through transcript in the *prsI/arsI* intergenic region, was likely to be transcribed in an operon with *prsI* (Fig. 3.10). *arsI* encodes a putative anti-anti-sigma factor, which due to its positioning is likely to regulate PrsI. The predicted protein product of the *arsI* gene has significant homology to SpoIIAA of *B. subtilis*, the anti-sigma factor antagonist which, with the anti-sigma factor SpoIIAB, regulates the activity of the sporulation-specific sigma factor σ^F (Yudkin and Clarkson, 2005). The homology included the conserved serine residue phosphorylated in SpoIIAA, determined using the NCBI Conserved Domain Database (http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml). Within the sequence upstream of *arsI* was an excellent match to the BldD binding consensus sequence AGTgA (n_m) TCACc (Elliot *et al.*, 2001): the *arsI* sequence was AGTGC (12) TCACG (Table 3.3, Fig. 3.20). The presence of potential binding

Figure 3.20

Sequence of the *arsI* coding region and surrounding area. Only the coding strand is shown. Start and stop codons are marked in bold italics, coding regions (*arsI* and part of the upsream *sigI*) are shown in bold, and the transcription start points are marked in italics with a bent arrow above (p_{arsI} and p_{sigI}). The putative BldD binding consensus sequence is marked in red, with a red arrow above each half site to indicate the first (forward arrow) or second (reverse arrow) portion of the half site. Primers are marked with an arrow indicating the strand they are derived from; forward primers are marked above the sequence, and reverse primers are indicated below. Wavy lines at the end of some primers indicate engineered restriction sites; the colour of the restriction enzyme is matched to the colour of the primer. Where primers include sequence from two lines in the figure, the arrow and primer name are colour coded.

*Bgl*II/*Eco*RI
CGA21/CGA63/CFU3

ggagggagtg tcgacgtcgc ttggtgggta cgcgatccgt cgagccgggg

tga*cat*gatg tcctccatcg ttctcggcat acggctgccg aagccaaata

cgtgcactgc ggtgtgcggc gcctccaaag ccggccgtgt cggtttacgt

gt*c*ttactag ccctaccggg tttaccgagc tgaccgcaag tgtgttctgt
P_{sigI}

gcgggtatgt ccgtttgtgt gcggttgttc gggtgtcgga gtgcggtggg

P_{arsI}
CGA47
aaggcgtagt gttcga*g*tgc gtcagcaagc agtcacgacc tcgggagaga
CGA62

gaaacggc*at g*gaccgcggg acggtcggca gtgcccagtc gggccggctt
*Xba*I
CGA48/CFU2 *Bgl*II

ctggtcgagg tacgggaaga ggggcccagc gccgtcgtga ccccggcggg
CGA19

tgagctggat caccacaccg ccgatctgtt gcgcgagccg ctggaggact

gtctcgacaa gggattcaac cggctcgtcg tggactgctc acggctggag

ttctgcgact ccacggggct caacgtcctg ctcggcgctc ggttgaaagc

*Bam*HI CGA20
ggaagccgcg gggggcggag tgcatttggt ggccatgcag ccggtagtgg

cgcgcgtgtt cgagatcacg gggccgaag cggtcttcac ccttcatgac

acgcttgcgg ccgccctggc cgacgcgtcc gac*tga*gccg tccttgtccg

CGA44
gccgacggag gggtgcgtgt gcttcccgga gcgggtgttt tcgccccgcg

ccgggtttcg tgcgcaatac aggggtgttc gtgcgatccg gtcgggcagg sites upstream of both genes, *arsI* and *prsI*, in a putative operon suggested that both the *prsI*-specific transcripts and the transcript containing both *prsI* and *arsI* were similarly regulated by BldD. The excellent consensus match in itself was sufficient to encourage further study of *arsI* as a genuine BldD binding target.

Interestingly, the microarray data did not support *arsI* being a BldD target. In the single experiment that was performed, induction of *bldD* had no effect on *arsI* transcription, even when compared to the control strain (without inducible *bldD*) (Fig. 3.9, C and D, Table 3.2). However, this in itself was not conclusive since two known repressed targets, *bdtA* and *whiG*, showed no repression upon *bldD* induction on the same array. This confirmed the need for further investigation before any conclusion could be drawn about whether or not *arsI* is a BldD target.

3.4.2 Analysis of *arsI* transcription patterns

As with *prsI*, transcript mapping of *arsI* was undertaken through both S1 nuclease mapping and primer extension experiments using RNA from solid culture time courses grown on rich medium and harvested at various points in development. Determination of the transcription start point was solely from high resolution S1 nuclease mapping (Fig. 3.21 A), but confirmation of BldD regulation of *arsI* expression also came from primer extension in the *prsI*/*arsI* intergenic region (Fig. 3.21 B). The 353 bp probe used for S1 nuclease mapping (CGA21-19) extended from 291 bp upstream of the *arsI* start codon to 57 bp inside the *arsI* coding region, including 43 bp inside the coding region of the divergently expressed gene *sigI*, and had an *Eco*RI site derived from CGA21 to provide the

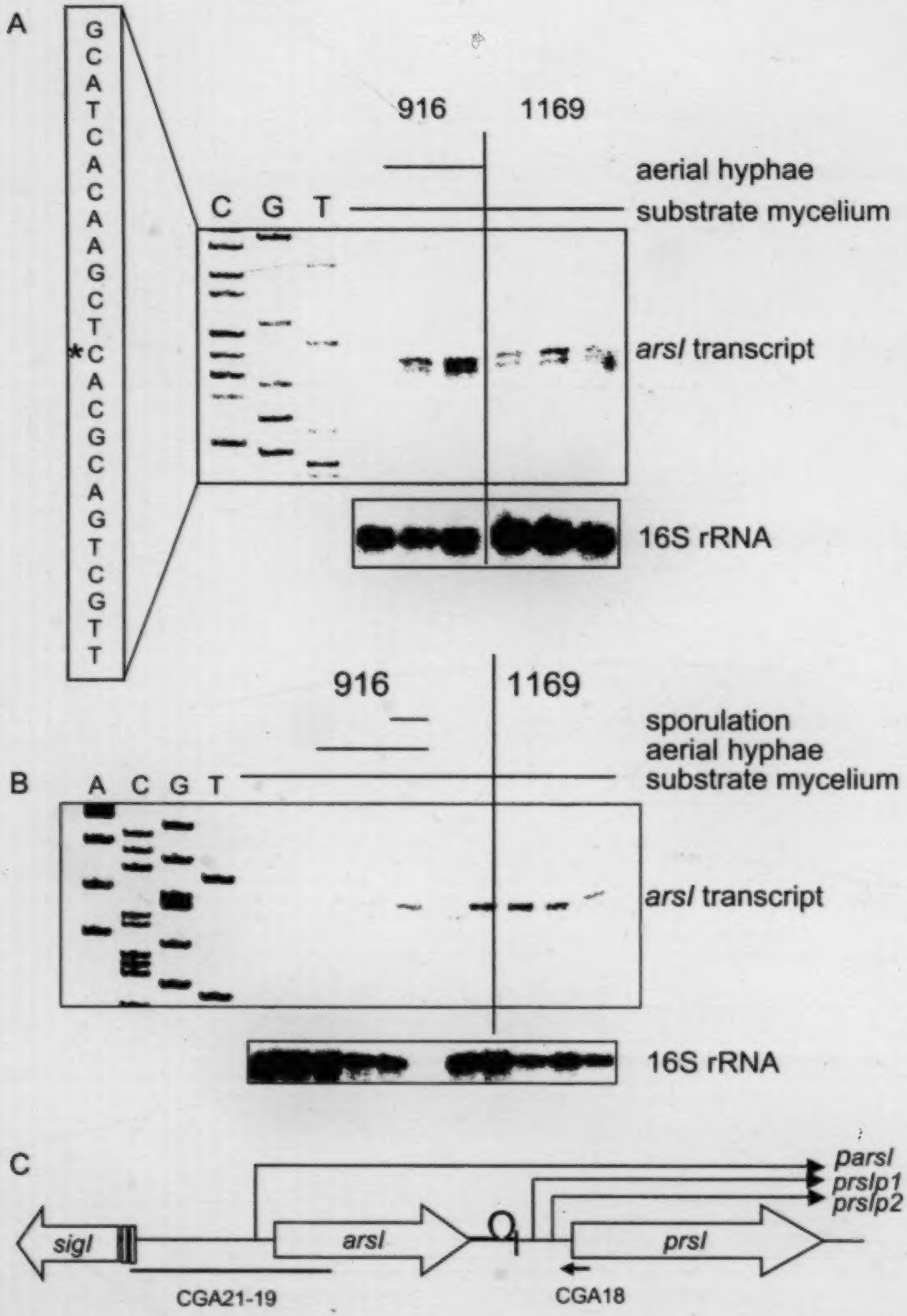
Figure 3.21

BldD regulation of *arsI* transcription shown by S1 nuclease mapping and primer extension.

A) High resolution S1 nuclease mapping of the *arsI* transcription start point using a 353 bp probe (CGA21-19). RNA was harvested at various points in development (indicated across the top) from surface-grown 916 (wild-type) and 1169 (*bldD* point mutant) cultures and hybridized with end-labelled probe overnight. A sequencing ladder (marked C G T) was generated using the primer from the protected end of the probe (CGA19). The ladder and mapping reactions were resolved on a 6% polyacrylamide sequencing gel. The transcription start point is indicated by an asterisk (*) in the sequence to the left of the gel. No internal control was included, so the same volumes of RNA were analyzed by northern hybridization, probed for 16S rRNA, and shown underneath as a control for the amount of RNA in each lane.

B) Primer extension to analyze expression from the *arsI* promoter in RNA isolated from wild type (916) and *bldD* mutant (1169) *S. coelicolor* strains over a time course of development. The primer used was CGA18, which overlaps the *prsI* start codon; the band shown is from p_{arsI}, in the intergenic region between *arsI* and *prsI* upstream of both of the *prsI*-specific promoters and downstream of a predicted stem-loop structure in the bicistronic transcript. The stage of development is indicated by the horizontal lines above the bands. The same volumes of RNA were analyzed by Northern hybridization using a probe for 16S rRNA (BKL54) to control for the amount of RNA in each lane.

C) Schematic illustrating the transcript quantification methods. High resolution S1 nuclease mapping was performed using the probe CGA21-19, and the *arsI* transcription start point identified. Primer extension was performed using CGA18, which overlaps the *prsI* start codon. The predicted stem-loop structure is identified by a loop in the *arsI-prsI* intergenic region, and the fall-off point used to quantify the *arsI* read-through transcript is identified by a vertical line immediately downstream of the predicted stem-loop structure.


A
G C A T C A C A A G C T *C A C G C A G T C G T T
916
1169
aerial hyphae
C G T
substrate mycelium
arsI transcript
16S rRNA
B
916
1169
sporulation
aerial hyphae
A C G T
substrate mycelium
arsI transcript
16S rRNA
C
parsI
prsIp1
prsIp2
sigI
arsI
prsI
CGA21-19
CGA18

necessary non-homologous end. The transcription start point of *arsI* was determined to be 42 bp upstream of the *arsI* translational start codon. Developmental regulation was clear in the wild type strain, with no transcript being produced until the appearance of aerial hyphae. This regulation was abolished in the *bldD* mutant strain where transcript was visible in all time points tested, indicating *bldD*-dependent regulation of transcription from p_{arsI}. Although S1 nuclease protection was only performed successfully on a full time course of RNA once, the results were confirmed by primer extension experiments on a full time course of RNA, and by S1 nuclease protection experiments on single time points from three different time courses. Primer extension experiments used the primer CGA18 (Fig. 3.11, Table 2.4), which anneals to a region overlapping the *prsI* start codon. Attempts to use CGA19, which anneals inside the *arsI* coding region (Fig. 3.20), were unsuccessful after several attempts. The largest band seen in the primer extension was a result of partial extension of the longer polycistronic transcript directed from p_{arsI} rather than from a third *prsI* promoter, and as such was used to confirm the timing of *arsI* expression. Analysis of the region between the stop codon of *arsI* and the partial extension product 101 bp downstream revealed a stem-loop structure (likely resulting in mRNA stabilization and blocking reverse transcription) with a ΔG of -49.75, supporting the identification of this band as being produced from the polycistronic transcript. Although, in contrast to the S1 nuclease protection results, this transcript seems to appear only after the onset of aerial hyphae formation, this could be explained by partial degradation of the transcript from the 3' end. Techniques that measure near the transcription start point would detect all transcription arising from the promoter, while techniques that measure considerably further down (such as primer extension within an operon) might show lower transcript levels due to transcript instability (3' degradation). The S1 nuclease mapping results were considered more reliable as they measured transcript levels at the transcription start point. Developmental regulation of p_{arsI} was confirmed on three independent time courses: one with S1 nuclease protection, and two with primer extension.

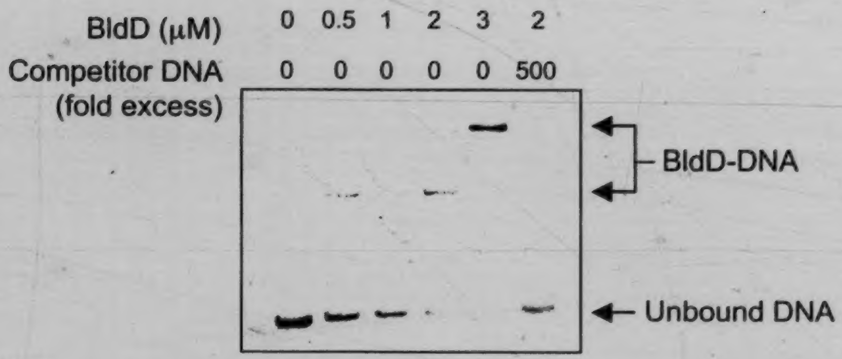
3.4.3 Binding of BldD to the *arsI* promoter region

The transcription mapping experiments that were performed suggested BldD regulation of *arsI*, but did not determine whether or not this regulation was through direct binding of BldD to the *arsI* promoter. Therefore, EMSA assays were conducted using increasing amounts of His_6-BldD (Elliot and Leskiw, 1999), which was incubated with 1 ng of labelled probe in the presence of poly d(I-C), then loaded on a running 8% glycerol-containing polyacrylamide gel. The purpose of the electrophoresis was to separate bound from unbound DNA. Experiments using a 69 bp probe (CGA47-48) confirmed BldD binding to the region upstream of *arsI* (Fig. 3.22). This small probe extended from -23 to +45 relative to the transcription start point of *arsI*. A clear shift was seen with as little as 0.5 μM BldD, and maximum shift was reached at 3 μM BldD. The K_D was 1.2 μM, higher than BldD binding to its own promoter (0.4 μM). Binding was significantly reduced by the addition of a 500-fold excess of unlabelled probe,

Figure 3.22

BldD binds to the *arsI* promoter region.

EMSA analysis of BldD binding to the *arsI* promoter region was performed using a 69 bp probe (CGA47-48). One nanogram of probe was end-labelled and incubated with increasing amounts of purified His$_6$-BldD, then loaded on an 8% polyacrylamide gel containing 1.5% glycerol to resolve the bound and unbound DNA. Competitor DNA was unlabelled probe added in 500-fold excess to confirm the specificity of binding. Bound and unbound DNA are indicated by arrows.


BldD (μM) 0 0.5 1 2 3 2
Competitor DNA (fold excess) 0 0 0 0 0 500
BldD-DNA
Unbound DNA

indicating that the binding was due to specific interactions of BldD with the *arsI* promoter. Results were confirmed with multiple experiments.

3.4.4 Determination of the site of BldD binding in the *arsI* promoter by DNase I footprinting

For the footprinting assays, probes were redissolved to give a constant number of picomoles of labelled probe. The first several attempts used a 353 bp probe (CGA21-19) - somewhat longer than the 257 bp probe used successfully to footprint BldD binding to the *bldD* promoter. CGA21-19 extends from 43 bp within the *sigI* coding region to 57 bp inside the *arsI* coding region. Consistently, a strong probe degradation band was seen when the footprinting reactions were resolved by electrophoresis, just as a degradation band had been seen with the 162 bp *prsI* probe CGA43-42 (Fig. 3.19). The same band was seen when the labelled probe was run by itself and without digestion with DNase I. In an attempt to eliminate this problem, the labelling of the primer and amplification of the probe were performed on the same day as the footprinting, but this did not reduce the intensity of the degradation band. The putative degradation band was not seen when unlabelled probe was resolved on a 6% polyacrylamide gel and stained with ethidium bromide, suggesting it was not a PCR artifact. This suggested that the degradation was a result of PCR amplification using a labelled oligonucleotide and that radiolysis was a possible explanation.

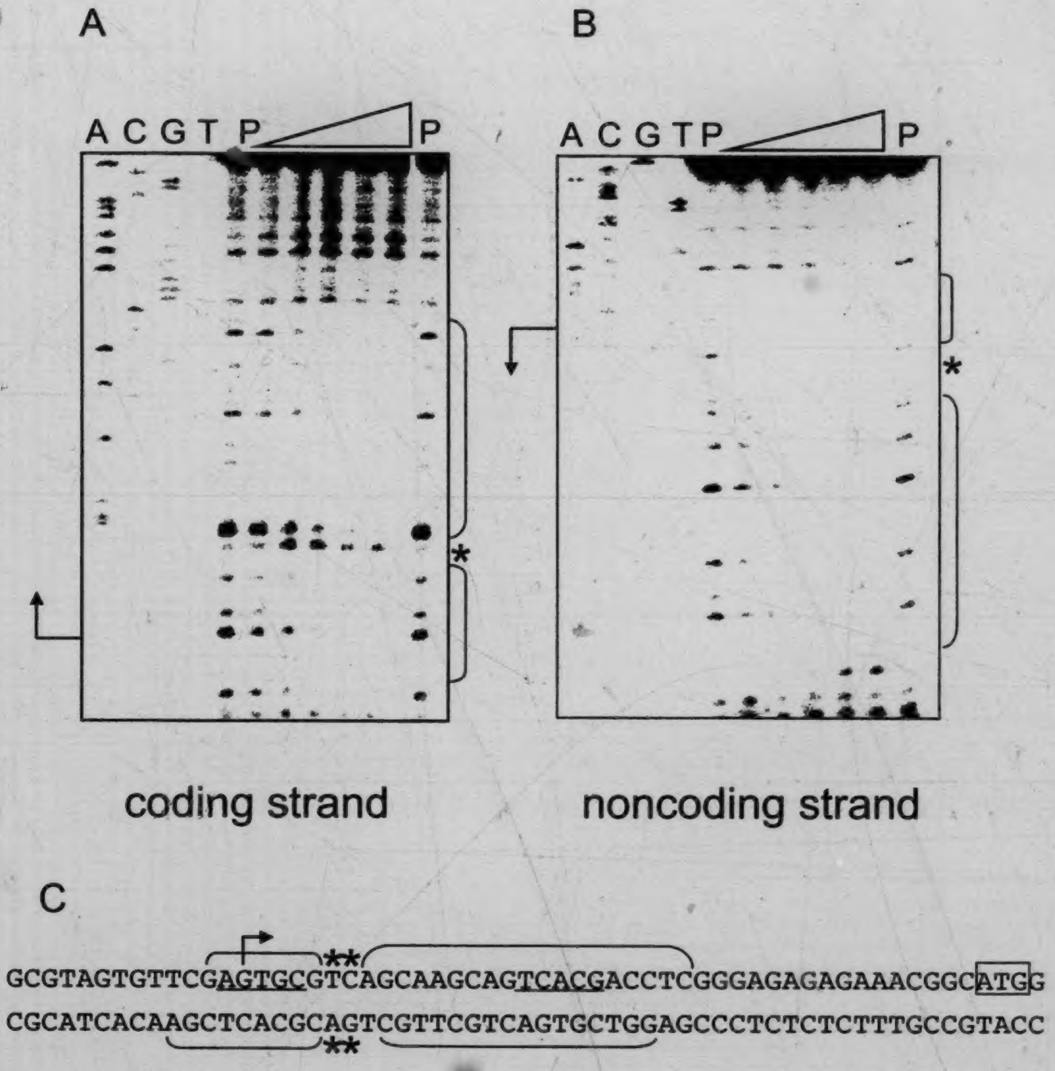
In an attempt to eliminate the degradation band, the PCR amplified, radiolabelled probe was denatured and purified directly from a 6% sequencing gel. The advantage of this method of purification as compared to purification from a 6% non-sequencing polyacrylamide gel was the much higher resolution of the full-length and degradation bands. Elution time was reduced from the standard overnight incubation to two or three hours, to minimize further degradation of the full-length band. When this purified fragment was run on a sequencing gel, the amount of degradation was reduced, not eliminated (not shown).

Digestion of probe that had been end-labelled after amplification was then used to try to circumvent the need for PCR amplification using a radiolabelled oligonucleotide. End-labelling of the probe, followed by *Eco*RI digestion (of the non-homologous tail on CGA21) yielded incomplete digestion, even after addition of up to 50 units of enzyme over two hours for nanogram amounts of DNA. Since attempts to use the incompletely digested probe in footprinting experiments failed to allow visualization of BldD binding, a different probe was designed.

After many attempts to footprint *arsl* using the long probe, a different, shorter probe was used. The same 69 bp CGA47-48 probe had been used successfully for EMSA (Fig. 3.22). Because of the short length of the probe, in this case the footprinting reactions were resolved using a 10% acrylamide gel (in contrast to the standard 6% used for the transcript mapping) to allow resolution of the small fragments. Clear footprints were seen on both strands overlapping the putative BldD binding consensus sequence (Fig. 3.23). The footprints define the two half sites of the sequence matching the consensus BldD binding site. As well as the footprints, a DNase I hypersensitive site can be seen on both strands

Figure 3.23

DNase I footprinting of the *arsI* promoter region using a 69 bp probe (CGA47-48). One primer was labelled before PCR, then the product purified from a polyacrylamide gel and precipitated with ethanol and glycogen. The purified probe was redissolved to allow 0.06 pmol per lane to be incubated with increasing amounts of purified His_6-BldD (0.5, 1, 2, 3 and 4 μM). Samples were then subjected to DNase I digestion and run on a 10% polyacrylamide sequencing gel to resolve the digestion products. A sequencing ladder (A C T G) generated from the primer that was labelled for the PCR was run beside the reactions to allow determination of the location of the footprint. P indicates lanes with probe alone (no BldD). The transcription start point is marked by a bent arrow, and footprints are indicated by brackets beside the gel picture. Footprinting of the coding strand is found in panel A and the non-coding strand in panel B. The sequence protected by BldD can be found at the bottom of the figure (C). Again, the footprint is marked with brackets and the transcription start point is indicated by a bent arrow. The start of the coding region (ATG) is boxed. The DNase I hypersensitive sites are marked in both the image and the sequence with an asterisk (*).



coding strand

noncoding strand

GCGTAGTGTTCGAGTGCGTCAGCAAGCAGTCACGACCTCGGGAGAGAGAAACGGCATGG
CGCATCACAAGCTCACGCAGTCGTTCGTCAGTGCTGGAGCCCTCTCTCTTTGCCGTACC

between the two halves of the BldD binding sequence, possibly indicating a bend in the DNA. Reactions were run next to a sequencing ladder generated with the labelled primer (CGA47 for the coding strand and CGA48 for the non-coding strand).

3.5 *sigI* is not a BldD target

3.5.1 Sequence analysis of the *sigI* promoter region

A third gene is encoded in the vicinity of *arsI* and *prsI*, divergent from the other two (Fig. 3.10). *sigI* (*sce25.09*, *sco3068*) encodes a sigma factor shown to be responsive to osmotic shock (Viollier *et al.*, 2003). SigI levels increased transiently upon addition of 0.5 M NaCl or 0.5 M sucrose in a transcriptionally-dependent manner (Viollier *et al.*, 2003). No BldD binding consensus sequence was seen upstream of the coding region and thus *sigI* is unlikely to be a direct target of BldD. As shown in sections 3.3.2 and 3.4.2, BldD regulates the transcription of *prsI* and *arsI*, which encode an anti-sigma factor and an anti-anti-sigma factor respectively. The products of these genes are likely to regulate the activity of the sigma factor encoded by *sigI*. Through regulation of *arsI* and *prsI*, BldD is therefore likely to indirectly affect the activity of SigI.

3.5.2 Analysis of *sigI* transcription patterns

To determine if *sigI* transcription is *bldD*-dependent, high resolution S1 nuclease mapping was carried out using the 312 bp probe CGA 62-63 which extends from 43 bp inside the *sigI* coding region to just inside the *arsI* coding region. The transcription start point of *sigI* was located 97 bp upstream of the start codon and 113 bp away from the divergent *arsI* transcription start point (Fig. 3.24). Expression of *sigI* was constitutive in both wild type and *bldD* mutant strains, suggesting that BldD does not regulate *sigI* transcription. Results were confirmed on an independent time course. A probe for *hrdB*, the principle sigma factor of *S. coelicolor* was included in the reactions as a loading control. These results are consistent with previously published data (Viollier *et al.*, 2003), showing no change in SigI protein levels when a *bldD* mutant and wild type strain were compared. A summary of the *sigI* locus, with the transcription start points marked, is found in Figure 3.15.

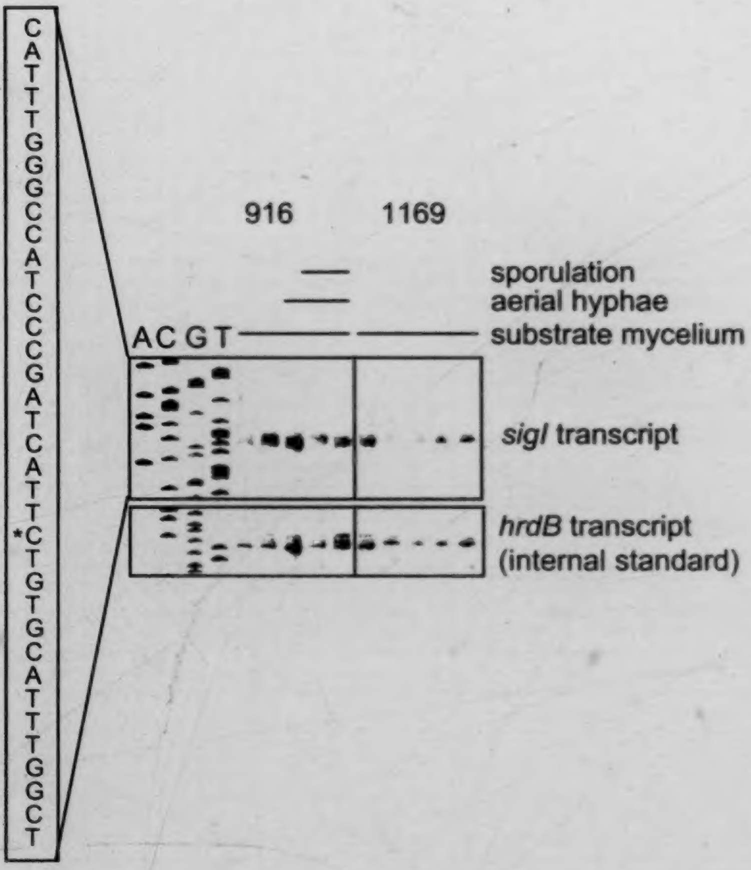
3.6 Creation of *prsI*, *arsI*, and *sigI* null mutants

3.6.1 PCR-targeted mutagenesis of *prsI*

As *prsI* is a BldD target and encodes a regulatory protein, there was a possibility that *prsI* itself would be a global regulator of differentiation or antibiotic production. However, the possibility also existed that it would exhibit no obvious phenotype, as important processes in *S. coelicolor* are often controlled by multiple pathways (Chater, 1998). Numerous attempts were made to delete *prsI* in the chromosome using the method of Gust et al (2003) (see section 2.7). This work was done by Linda Bui as a summer student under the direction of C. Galibois. A schematic of the procedure can be found in Figure 2.2. The apramycin cassette was amplified with the primers prsIdelF and prsIdelR (Table 2.4), which each contain a 39 bp extension matching sequence flanking the *prsI* gene to facilitate recombination and replacement of the wild type *prsI* gene with

Figure 3.24

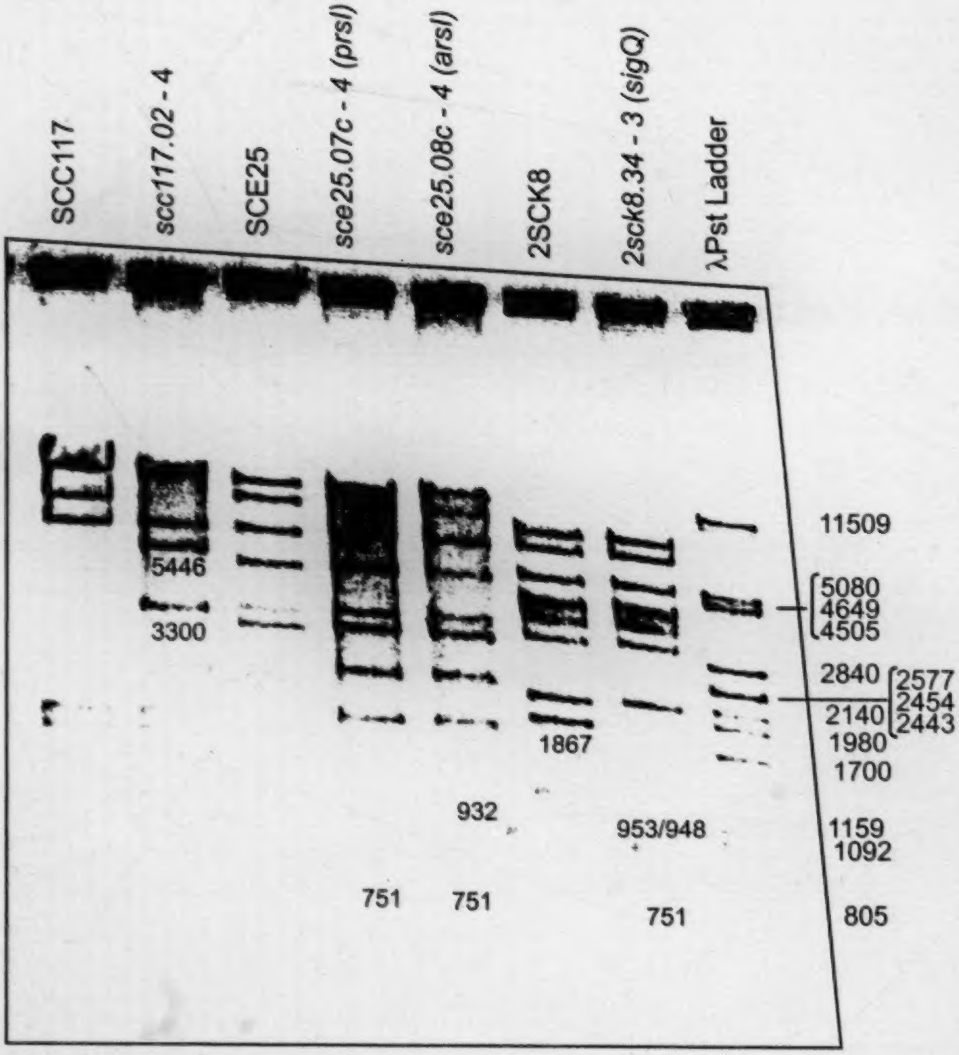
High resolution S1 nuclease mapping of the *sigI* transcription start point using a 312 bp probe (CGA62-63). RNA was harvested at various points in development (indicated across the top) from surface-grown *S. coelicolor* 916 (wild-type) and 1169 (*bldD* point mutant) cultures, and hybridized overnight with the end-labelled probe. A sequencing ladder (marked A C G T) was generated using the primer from the protected end of the probe (CGA63). The ladder and mapping reactions were resolved on a 6% polyacrylamide sequencing gel. The transcription start point indicated by an asterisk (*) in the sequence to the left of the gel. An internal standard to control for RNA loading is provided by the 110 bp *hrdB* probe (hrdB-S1 - hrdB-R).


CATTTGGGCCATCCCGATCATT*CTGTGCATTTGGCT
916
1169
sporulation
aerial hyphae
A C G T
substrate mycelium
sigI transcript
hrdB transcript
(internal standard)

the apramycin cassette. The disruption cassette was then transformed into *E. coli* strain BW25113/pIJ790 containing cosmid SCE25 (Redenbach *et al.*, 1996) which had been grown with arabinose to induce the λ-RED system. Replacement of the wild type gene in cosmid SCE25 with the disruption cassette was confirmed by digestion of isolated cosmid DNA with *Eco*RI and *Sac*I and separation of the digestion products by electrophoresis on a 1% agarose gel (Fig. 3.25). The disrupted cosmid was transferred into *E. coli* strain ET12567/pUZ8002 and then into *S. coelicolor* strain M600 by conjugation. Fifty exconjugants were initially transferred to R2YE plates, and then inoculated onto DNA agar with nalidixic acid and apramycin, and with or without kanamycin. The nalidixic acid was used to prevent growth of any residual *E. coli*, the apramycin selected for the disruption cassette, and kanamycin sensitivity was used to screen for the second crossover event that would result in loss of the cosmid and one copy of the gene. Since no kanamycin sensitive isolates were found among the first 50 exconjugants tested, fifty more primary exconjugants were selected. Again, only kanamycin resistant colonies were seen. A small-scale mycelial stock was made of one colony and inoculated on multiple non-selective plates to generate approximately 4000 colonies. These were then screened for apramycin resistance paired with kanamycin sensitivity. Some kanamycin sensitive isolates were found, but since these were also apramycin sensitive, they were considered to be wild type revertants. No apramycin resistant, kanamycin sensitive colonies were isolated. Therefore, *prsI* appeared to be an essential gene.

Figure 3.25

PCR-targeted mutagenesis of four *S. coelicolor* genes. PCR-targeted mutagenesis of four genes [*scc117.02* (*sco2529*, see Appendix D), *prsI* (*sce25.07c/sco3066*), *arsI* (*sce25.08c/sco3067*) and *sigQ* (*2sck8.34/sco4908*)] was performed to confirm replacement of the wild type genes in their respective cosmids with the apramycin resistance cassette. Cosmid DNA was digested with *Eco*RI and *Sac*I, enzymes that produce a different pattern of restriction fragments in the wild type and disrupted cosmids. Digestion products were separated by electrophoresis on a 1% agarose gel. Sizes of fragments present only in the wild type (designated with uppercase letters) or disrupted cosmids are indicated next to the appropriate bands. The marker used was lambda phage DNA digested with *Pst*I.

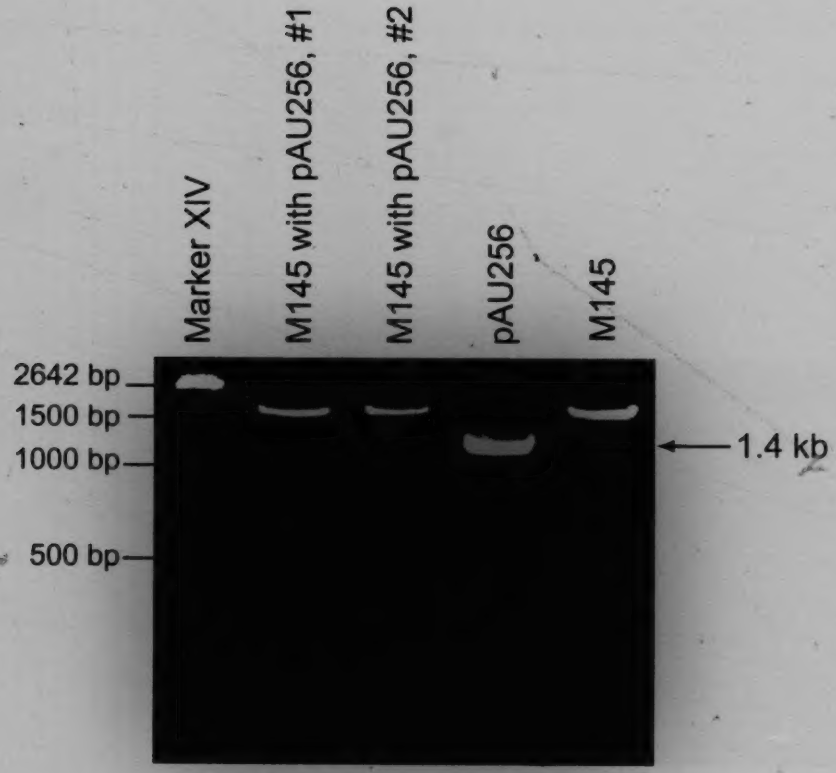

SCC117
scc117.02 - 4
SCE25
sce25.07c - 4 (prsI)
sce25.08c - 4 (arsI)
2SCK8
2sck8.34 - 3 (sigQ)
λPst Ladder
5446
3300
1867
932
953/948
751
751
751
11509
5080
4649
4505
2840
2577
2454
2443
2140
1980
1700
1159
1092
805

Since *prsI* appeared to be essential, later disruption attempts focused on disruption of the wild type copy of *prsI* in the presence of a second chromosomally integrated (in the ϕC31 *att* site) copy of the gene. This work was performed by Kent Gislason, as a technician under the direction of C. Galibois. The primers CGA21 and CGA65 were used to amplify the coding regions of both *prsI* and *arsI* from cosmid SCE25 (Redenbach *et al.*, 1996); the primers amplify a region from 41 bp inside the *sigI* coding region to 60 bp downstream of the *prsI* coding region. The reason for including both genes was that the transcript that begins upstream of *arsI* includes *prsI*, and full complementation might not be possible without all of the native promoters. The amplified fragment was then cloned into pCR®2.1 TOPO to generate pAU253.

pAU253 was digested with *Eco*RI to liberate the CGA21-65 insert from the pCR®2.1 TOPO multiple cloning site. The resulting fragment was ligated into *Eco*RI-digested pSETΩ (O'Connor *et al.*, 2002) to generate pAU256. This plasmid was conjugated into wild type *S. coelicolor* (M145) via the non-methylating *E. coli* strain ET12567/pUZ8002. Two spectinomycin resistant exconjugants were isolated, and chromosomal DNA was isolated from both. Primers JST7 and JST8, which flank the pSETΩ polylinker, were used in the subsequent PCR amplification reactions. Chromosomal DNA from the wild type parent strain M145 was used as a negative control, and pAU256 was used as a positive control (Fig. 3.26). Roche Molecular Weight Marker XIV was used to determine band sizes. The expected 1.4 kb fragment was amplified from the plasmid but not from the spectinomycin resistant strains, indicating spontaneous

Figure 3.26

Confirmation of spontaneous spectinomycin resistance in pAU256 exconjugants. Chromosomal DNA was isolated from M145 (wild type) and two strains displaying spectinomycin resistance following conjugation to introduce pAU256 (pSETΩ containing the full *arsI* and *prsI* coding regions). Primers JST7 and JST8, which flank the pSETΩ polylinker, were used in an attempt to amplify the expected 1.4 kb product. The original pAU256 plasmid DNA was used as a positive control. The 1.4 kb band is indicated by an arrow. Roche Molecular Weight Marker XIV was used to determine band sizes.


Marker XIV
M145 with pAU256, #1
M145 with pAU256, #2
pAU256
M145
2642 bp
1500 bp
1000 bp
500 bp
1.4 kb

spectinomycin resistance rather than proper integration of the plasmid. The construct was deemed not to be useful for use in conjunction with PCR targeted mutagenesis. Without this vector, no second copy of *prsI* could be introduced and PCR-targeted mutagenesis of *prsI* was not possible. This project was therefore abandoned.

3.6.2 PCR-targeted mutagenesis of *arsI*

To help determine its role in *S. coelicolor* development, a null mutant of *arsI* was generated using PCR targeted mutagenesis (Gust *et al.*, 2003). This work was done by Linda Bui as a summer student under the direction of C. Galibois. The primers CGA58 and CGA59 were used to amplify the apramycin resistance disruption cassette. Recombination of the amplified cassette into cosmid SCE25 (Redenbach *et al.*, 1996) was confirmed by digestion with *Eco*RI and *Sac*I (Fig. 3.25) before transformation of the cosmid into the non-methylating *E. coli* strain ET12567 (pUZ8002) and introduction by conjugation into *S. coelicolor* strain M600. Thirty-four apramycin resistant exconjugants were selected, and two apramycin resistant, kanamycin sensitive isolates were selected for further analysis. Mycelial stocks were made and serially diluted for single colonies. Isolated colonies arising from the selected exconjugants had a bald phenotype. Chromosomal DNA was isolated from one kanamycin resistant and three kanamycin sensitive single colonies. The double crossover (replacement of the wild type gene with the apramycin cassette) was confirmed by Southern hybridization of *Eco*RI/*Bam*HI digested chromosomal DNA (Fig. 3.27). Three separate probes were used: the apramycin cassette, a

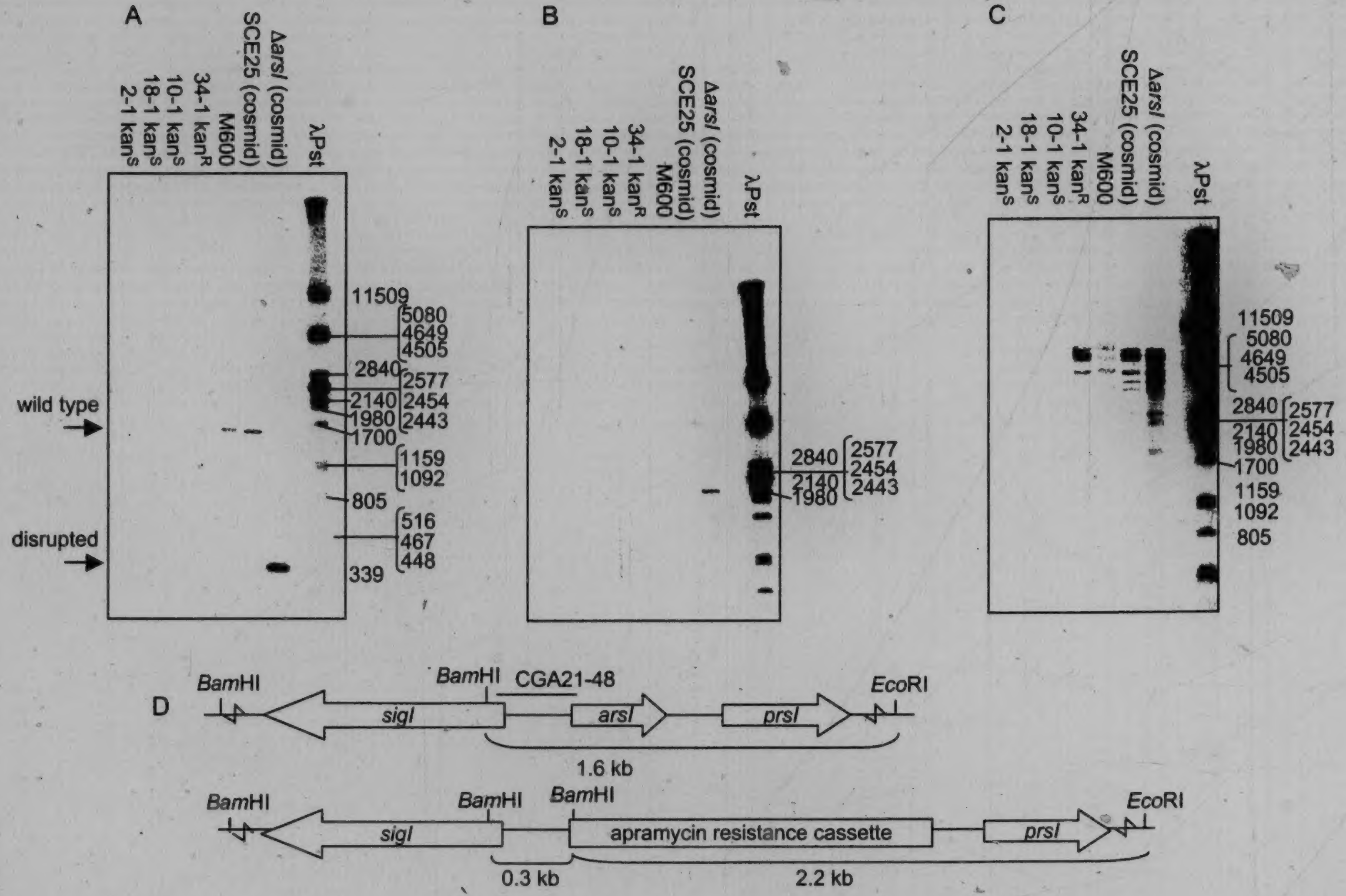
Figure 3.27

Confirmation of replacement of the wild type *arsI* gene with the apramycin resistance cassette in the chromosome of *S. coelicolor*. This work was performed by L. Bui under the direction of C. Galibois. Chromosomal DNA was isolated from mutants and digested with *Eco*RI and *Bam*HI before being separated by electrophoresis on an agarose gel, transferred to a nylon membrane, and hybridized with a radioactively-labelled probe (see below). The marker used in all cases was lambda phage DNA digested with *Pst*I (λPst).

A) The gene-specific probe used was a 299 bp PCR product (CGA21-48) spanning the intergenic region between *arsI* and *sigI*. Shown are the hybridizing wild-type (larger) and disrupted (smaller) fragments resulting from the restriction digest.

B) The apramycin cassette was used as a probe to confirm the presence of the cassette in the double (kanamycin sensitive) and single (kanamycin resistant) crossover mutants. Similarly digested chromosomal DNA from the wild type *S. coelicolor* (M600) and the undisrupted cosmid (SCE25) were included as negative controls. The disrupted cosmid (Δ*arsI*) was included as a positive control.

C) Wild type cosmid (SCE25) was used as a probe to confirm the absence of any gross chromosomal rearrangements caused by the recombination that resulted in replacement of *arsI* with the apramycin resistance disruption cassette.


A
B
C
λPst
Δ*arsI* (cosmid)
SCE25 (cosmid)
M600
34-1 kan^R
10-1 kan^S
18-1 kan^S
2-1 kan^S
wild type
disrupted
11509
5080
4649
4505
2840
2577
2454
2443
2140
1980
1700
1159
1092
805
516
467
448
339
D
*Bam*HI
*Bam*HI
CGA21-48
*Eco*RI
sigI
arsI
prsI
1.6 kb
*Bam*HI
*Bam*HI
*Bam*HI
*Eco*RI
sigI
apramycin resistance cassette
prsI
0.3 kb
2.2 kb

gene–specific probe, and the entire SCE25 cosmid (to test for gross chromosomal rearrangements). The gene specific probe was a 299 bp PCR product (CGA21-48) spanning the intergenic region between *arsI* and *sigI*. Use of the gene-specific probe showed a copy of the disrupted and wild type genes in the kanamycin resistant isolate but only the wild type gene in the undisrupted cosmid and wild type *S. coelicolor* chromosomal DNA. As expected, the disrupted cosmid and kanamycin sensitive strains had only the disrupted copy of the gene. Also as expected, the apramycin resistance cassette was seen in DNA isolated from both the kanamycin sensitive and kanamycin resistant exconjugants but not in chromosomal DNA from M600. The disrupted and wild type cosmids served as positive and negative controls and resulted in the expected pattern of hybridization. No gross chromosomal rearrangements in the vicinity of *arsI* resulted from replacement of *arsI* with the disruption cassette.

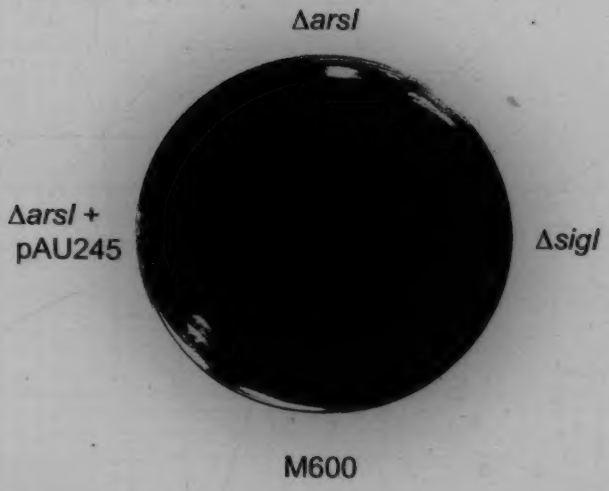
The *arsI* mutants were bald (lacking aerial hyphae) (Fig. 3.28) and had a slight reduction in blue pigment production, suggesting that *arsI* plays a role in the regulation of development in *S. coelicolor*. Alternately, the sigma factor target of *arsI* could direct the transcription of a repressor, the absence of which would result in unregulated expression of its target genes and thus a bald phenotype unrelated to the genes' function.

To confirm that the phenotype was the result of deletion of *arsI* and not caused by a secondary mutation, the strain was complemented. CGA21-CGA50 was amplified by PCR and digested with *Eco*RI to result in the formation of a fragment with one sticky end (*Eco*RI) and one blunt end. CGA21 (containing an

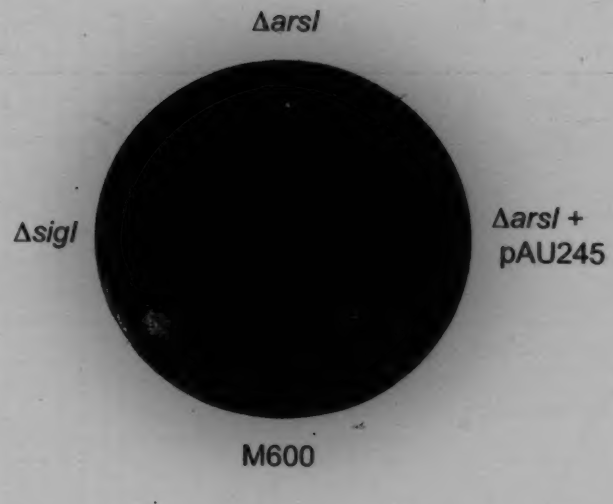
Figure 3.28

arsI, but not *sigI*, mutants have a bald phenotype. Generation of the null mutants was performed by L. Bui under the direction of C. Galibois. Null mutants of *arsI* and *sigI* are shown with the congenic wild type parent strain, *S. coelicolor* M600. The arsI null mutant was complemented with pAU245, containing the coding region of *arsI* and the entire intergenic region between *arsI* and the divergently encoded *sigI*, in the integrating plasmid pSETΩ. Strains were grown on R2YE for three days. Plates are shown in the bottom (A) and top (B) views.

A


Δ*arsI*
Δ*arsI* +
pAU245
Δ*sigI*
M600

B


Δ*arsI*
Δ*sigI*
Δ*arsI* +
pAU245
M600

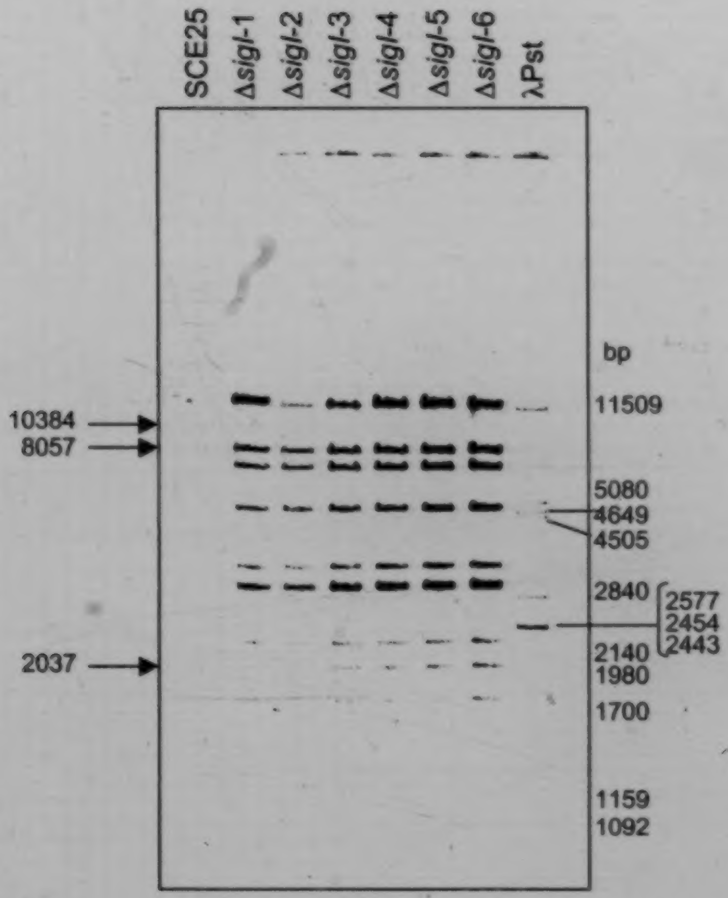
engineered *Eco*RI site at the 5' end) anneals inside the *sigI* coding region, and CGA50 anneals 14-33 bp downstream of the *prsI* coding region. pSETΩ (O'Connor *et al.*, 2002) was digested with *Eco*RI (sticky) and *Eco*RV (blunt) then ligated with the *Eco*RI-digested PCR product, creating pAU245. This integrating plasmid containing the full coding sequence of *arsI*, and all of its upstream sequence, was then conjugated into a representative *arsI* mutant designated Δ*arsI*. The resulting complemented strain was able to produce aerial mycelia at a near-wild type level (Fig. 3.28).

3.6.3 PCR-targeted mutagenesis of *sigI*

sigI was known to be involved in response to osmotic shock (Viollier *et al.*, 2003), but a disruption mutant had the potential to reveal other interesting things about its function. Again, Linda Bui, as a summer student under the direction of C. Galibois, generated the disruption using the method of Gust *et al.* (2003). The purified apramycin cassette was amplified with the sigI-delF and sigI-delR primers (Table 2.4), and recombination into SCE25 (Redenbach *et al.*, 1996) was confirmed by digestion with *Eco*RI and *Sac*I (Fig. 3.29). The disrupted cosmid was transferred to the *E. coli* strain ET12567 (pUZ8002) and then introduced by conjugation into the wild type *S. coelicolor* strain M600. One hundred exconjugants were screened as described above (section 3.6.1), and 23 kanamycin-sensitive strains were isolated. All kanamycin-sensitive colonies displayed a wild type phenotype. Eight colonies were chosen to make small-scale spore stocks, which were then plated to extinction, and single colonies were screened to confirm their developmental phenotype and kanamycin

Figure 3.29

Confirmation of replacement by PCR-targeted mutagenesis of the wild type *sigI* gene (*sce25.09/sco3065*) in cosmid SCE25 with the apramycin resistance cassette. This work was performed by L. Bui under the direction of C. Galibois. Cosmid DNA was digested with *Eco*RI and *Sac*I to produce a different pattern of restriction fragments in the wild type (SCE25) and disrupted cosmids (Δ*sigI*-1 to Δ*sigI*-6). Digestion products were separated by electrophoresis on a 1% agarose gel. Fragments (with sizes indicated) present only in the wild type or disrupted cosmids are marked by arrows. The size marker used was lambda phage DNA digested with *Pst*I (λPst).


SCE25
Δ*sigI*-1
Δ*sigI*-2
Δ*sigI*-3
Δ*sigI*-4
Δ*sigI*-5
Δ*sigI*-6
λPst
bp
10384
8057
2037
11509
5080
4649
4505
2840
2577
2454
2443
2140
1980
1700
1159
1092

sensitivity. Five of these single colonies were cultured for further analysis. Chromosomal DNA was isolated from the five kanamycin sensitive isolates and from one kanamycin resistant strain (for comparison). Southern analysis was performed on the *KpnI*/*Xba*I digested chromosomal DNA (Fig. 3.30). The same three probes were used for testing the *sigI* disruption as for the *arsI* disruption. Use of the gene-specific probe showed only the wild type gene in the wild type cosmid and *S. coelicolor* strain. The disrupted cosmid and kanamycin sensitive strains had only the disrupted copy of the gene. Although the kanamycin resistant strain had only the disrupted copy of the gene, this could have been from a false kanamycin resistance due to a secondary mutation. The apramycin resistance cassette was seen in all of the strains but M600, as expected, and in the disrupted cosmid but not the wild type cosmid. No gross chromosomal rearrangements were present in the vicinity of *sigI*. Strain 4-2-16-7a was chosen for further phenotypic analysis and designated Δ*sigI*. The disrupted strain had no visible impairment in aerial mycelium formation, antibiotic production, or sporulation (Fig. 3.28).

The product of *sigI*, σ^I, had been shown to be induced by osmotic shock; protein levels increased in the presence of 0.5 M NaCl or 0.5 M sucrose (Viollier *et al.*, 2003). Therefore, to see if disruption of *sigI* has any visible effect on the colony's ability to deal with hyperosmolarity (a concentration of salt or other solutes outside of the cell higher than that inside the cell), cultures were grown on minimal medium with glucose as the carbon source (MM glc) or MM glc with 0.5 M NaCl, 1 M NaCl or 1 M sucrose (Fig. 3.31). No difference was seen

Figure 3.30

Confirmation of replacement of the wild type *sigI* gene with the apramycin resistance cassette in the chromosome of *S. coelicolor*. This work was performed by L. Bui under the direction of C. Galibois. Chromosomal DNA was isolated from mutants and digested with *Kpn*I and *Xba*I before separation by electrophoresis on an agarose gel, transfer to a nylon membrane, and hybridization with a radioactively-labelled probe (see below). The marker used in all cases was lambda phage DNA digested with *Pst*I (λPst).

A) The gene-specific probe used was a 299 bp PCR product (CGA21-48) spanning the intergenic region between *arsI* and *sigI*. Shown are the hybridizing wild-type (larger) and disrupted (smaller) fragments resulting from the restriction digest.

B) The apramycin cassette was used as a probe to confirm the presence of the cassette in the double (kanamycin sensitive) and single (kanamycin resistant) crossover mutants. Similarly digested chromosomal DNA from the wild type *S. coelicolor* (M600) and the undisrupted cosmid (SCE25) were included as negative controls. The disrupted cosmid (Δ*sigI*) was included as a positive control.

C) Wild type cosmid (SCE25) was used as a probe to confirm the absence of any gross chromosomal rearrangements caused by the recombination that resulted in replacement of *sigI* with the apramycin resistance disruption cassette.

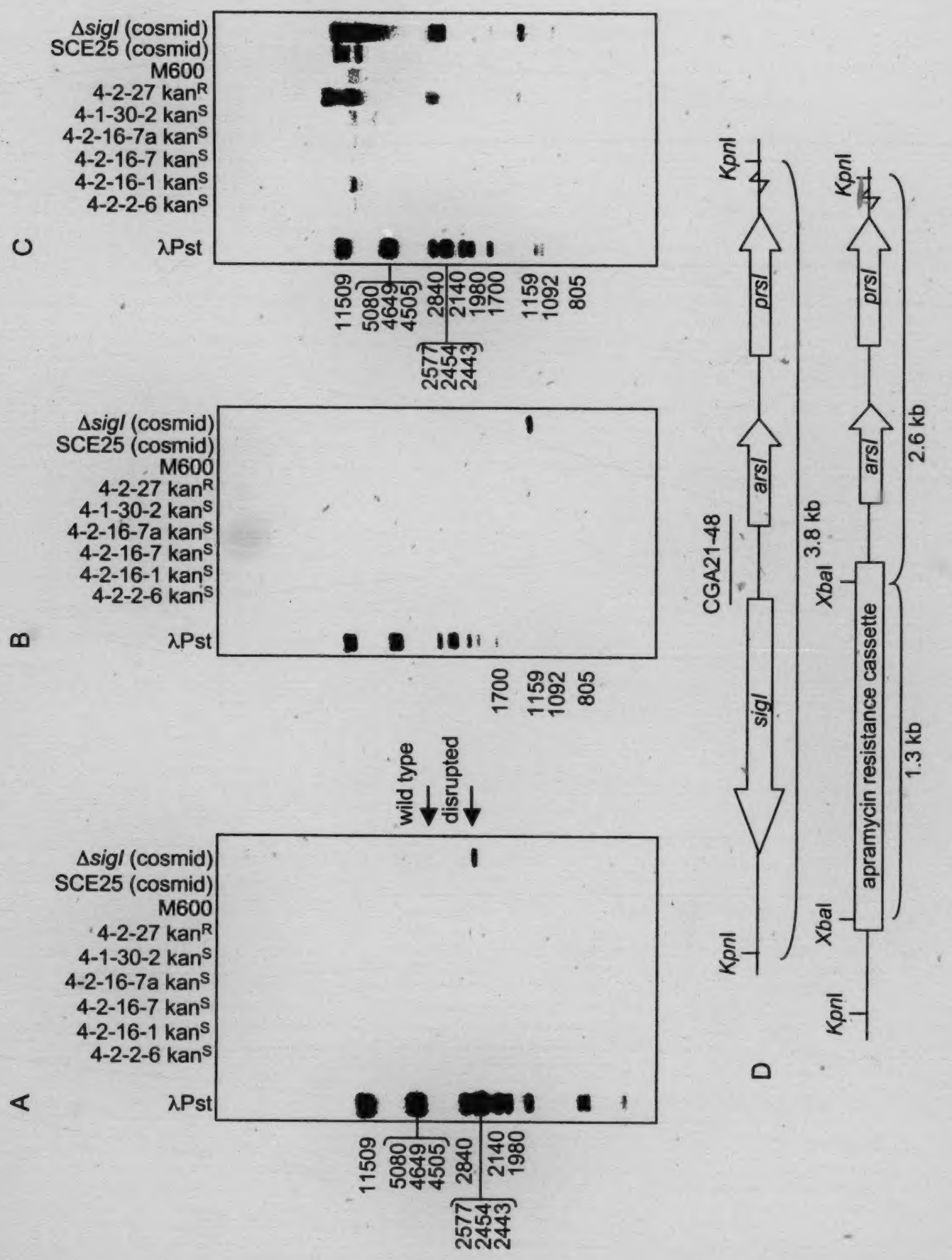

A
B
C
D
Δ*sigI* (cosmid)
SCE25 (cosmid)
M600
4-2-27 kan^R
4-1-30-2 kan^S
4-2-16-7a kan^S
4-2-16-7 kan^S
4-2-16-1 kan^S
4-2-2-6 kan^S
λPst
wild type
disrupted
11509
5080
4649
4505
2840
2577
2454
2443
2140
1980
1700
1159
1092
805
CGA21-48
KpnI
sigI
arsI
prsI
3.8 kb
XbaI
apramycin resistance cassette
1.3 kb
2.6 kb

Figure 3.31

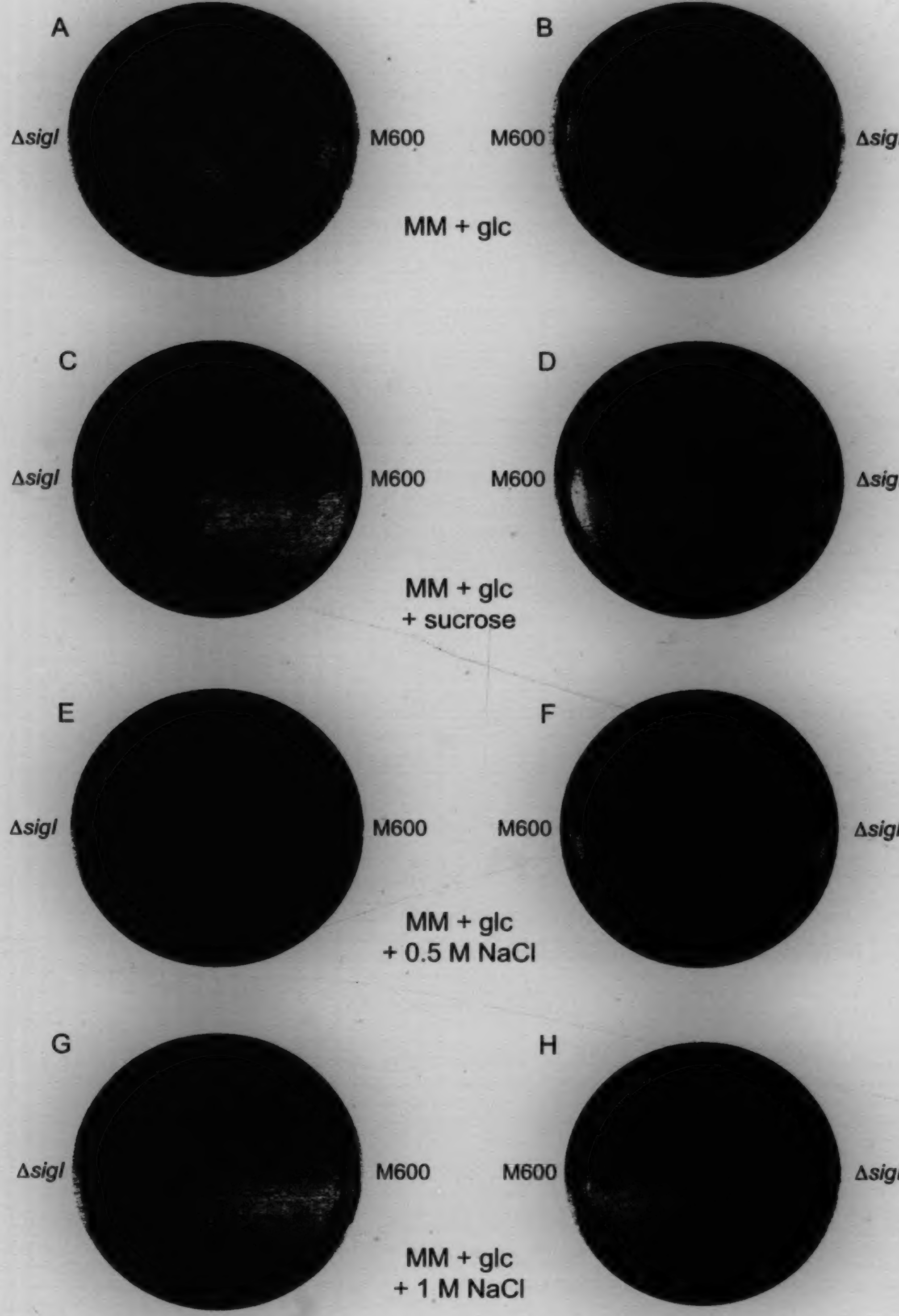
Δ*sigI* mutants show no change in response to osmotic shock. Osmotic shock testing was performed on the *sigI* mutants and compared to the parent strain M600. Strains were grown for 48 h on minimal medium with glucose (MM + glc) or the same medium supplemented with either NaCl or sucrose.

A and B) Front and back views, respectively, grown on MM + glc.

C and D) Front and back views, respectively, grown on MM + glc supplemented with 1 M sucrose (MM + glc + sucrose).

E and F) Front and back views, respectively, grown on MM + glc supplemented with 0.5 M NaCl (MM + glc + 0.5 M NaCl).

G and H) Front and back views, respectively, grown on MM + glc supplemented with 1 M NaCl (MM + glc + 1 M NaCl).


A
B
ΔsigI
M600
M600
ΔsigI
MM + glc
C
D
ΔsigI
M600
M600
ΔsigI
MM + glc
+ sucrose
E
F
ΔsigI
M600
M600
ΔsigI
MM + glc
+ 0.5 M NaCl
G
H
ΔsigI
M600
M600
ΔsigI
MM + glc
+ 1 M NaCl

between the parent strain (M600) and the *sigI* null mutant. During the course of this work, results were published from two separate labs that found a *sigI* null mutant to have a wild type phenotype (Lee *et al.*, 2005; Viollier *et al.*, 2003).

3.7 Analysis of cell specificity of transcription

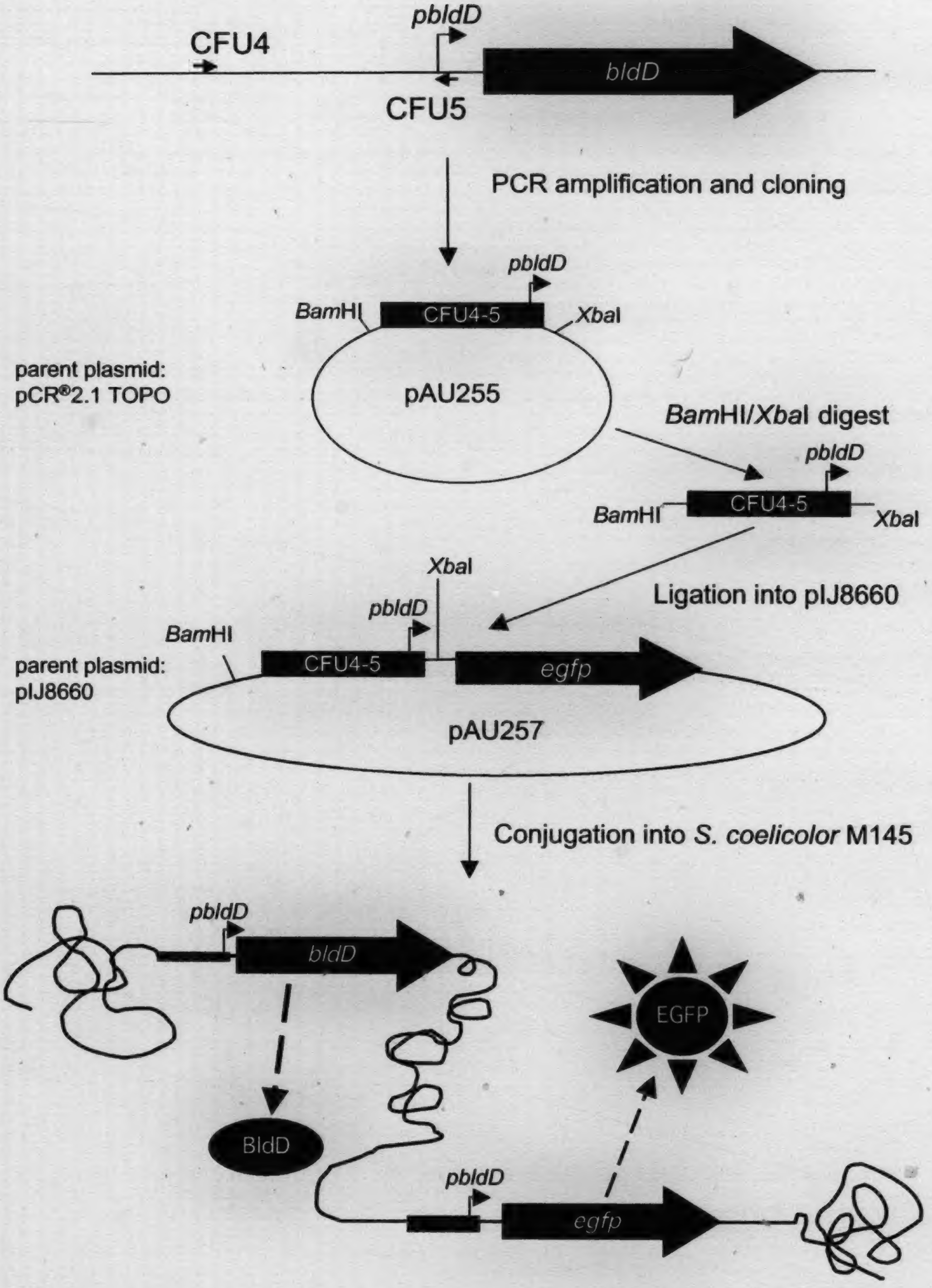
3.7.1 Examination of *bldD* expression within the colony

Spatial regulation of the transcription of the BldD target gene *sigH* had been demonstrated (Kelemen *et al.*, 2001), but the spatial expression of *bldD* itself had never been examined. To investigate this, a fragment containing the *bldD* promoter region was cloned upstream of a promoterless *egfp* gene and the resulting fluorescence examined in a wild type *S. coelicolor* strain. Initial cloning was carried out by C. Fung as a summer student under the direction of C. Galibois, and the work was completed by technician K. Gislason, again under the direction of C. Galibois. The primers CFU4 and CFU5 were used to amplify a 394 bp fragment upstream of *bldD* from cosmid SC9C5. The fragment extended from 336 bp upstream to 37 bp downstream of the *bldD* transcription start point. This fragment was then cloned into pCR®2.1 TOPO (Invitrogen) to create pAU255 (Fig. 3.32), and sequenced. The fragment was then transferred to pIJ8660 (Sun *et al.*, 1999) as a *Bam*HI-*Xba*I fragment, generating pAU257. Directional cloning ensured that the fragment was inserted in the correct orientation for *bldD*-driven *egfp* expression. pAU257 was passaged through the non-methylating *E. coli* strain ET12567 carrying the non-transmissible helper plasmid pUZ8002 and conjugated into the wild type *S. coelicolor* strain M145.

Figure 3.32

Construction of a *bldD-egfp* transcriptional fusion. This work was performed by C. Fung and K. Gislason under the direction of C. Galibois. First, the *bldD* promoter was amplified from cosmid SC9C5 using primers CFU4 and CFU5. The fragment was then inserted directly into pCR®2.1 Topo, generating pAU255. The fragment was then excised by digestion with *Bam*HI and *Xba*I to allow directional ligation into similarly digested pIJ8660, creating pAU257. pIJ8660 contains a promoterless *egfp* gene which will only be expressed in the presence of a suitable promoter fragment inserted upstream. The *bldD* promoter fragment is represented by a blue box, residual pCR®2.1 Topo sequence is indicated by black lines and pIJ8660 sequence is indicated by green lines or a green box.

The recombinant plasmid was introduced into the wild type *S. coelicolor* strain M145 by conjugation, where it integrated into the chromosome at the ϕC31 *attB* site.



pbldD
CFU4
bldD
CFU5
PCR amplification and cloning
pbldD
BamHI
CFU4-5
XbaI
parent plasmid:
pCR®2.1 TOPO
pAU255
BamHI/XbaI digest
pbldD
BamHI
CFU4-5
XbaI
XbaI
Ligation into pIJ8660
pbldD
BamHI
CFU4-5
egfp
parent plasmid:
pIJ8660
pAU257
Conjugation into S. coelicolor M145
pbldD
bldD
EGFP
BldD
pbldD
egfp

Integration of the pIJ8660 derivative at the ϕC31 *attB* site was confirmed by chromosomal PCR using the primers CFU4 and KTA-GFP-Rev (Bignell *et al.*, 2005). At the same time, as a control, pIJ8660 was conjugated separately into strain M145.

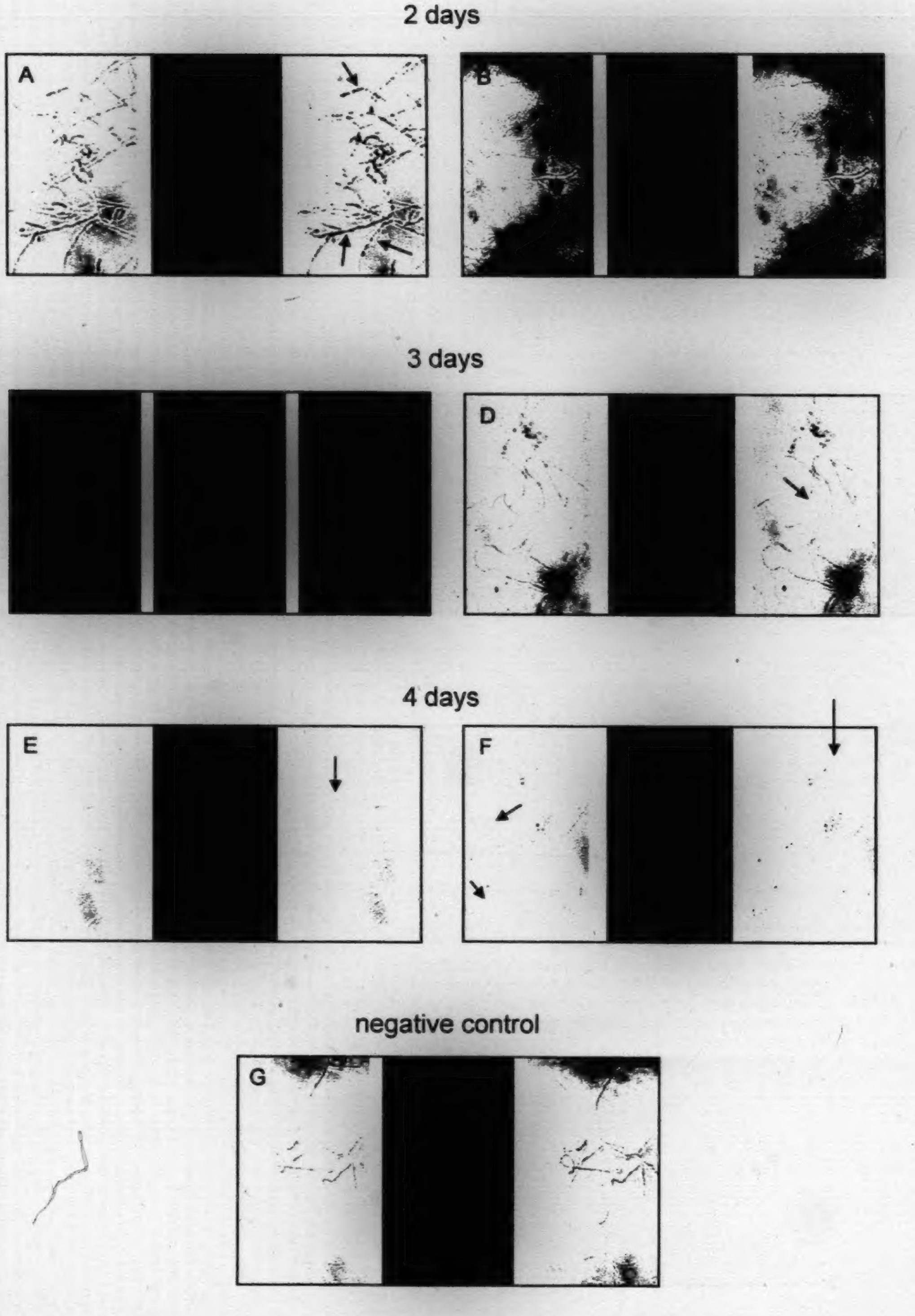
To examine the spatial transcription patterns, spores were used to inoculate R2YE solid medium at the intersection with sterilized glass cover slips (Gold Seal®) inserted into the medium at a 45° angle and grown for two, three or four days. *S. coelicolor* autofluoresces to a significant degree, so the control strain (M145 with pIJ8660) was used to adjust the detector to eliminate background fluorescence. Expression was examined through fluorescence and differential interference contrast images, and overlays of the two. After two and three days, *bldD* fluorescence was clearly seen throughout the colony, in all tissues from substrate mycelium to developed spores (Fig. 3.33) indicating universal expression of *bldD*. In contrast, however, no fluorescence above background was seen after four days of growth. The pattern of *bldD* transcription decreasing in later development corresponds well to the pattern of *bldD* transcription seen in the wild type *S. coelicolor* strain J1501, where no *bldD* expression is seen later in growth, although the transcription in this strain is undetectable after only 36 hours of growth (Elliot *et al.*, 1998).

3.7.2 Examination of spatial regulation of *prsI* transcription

Experiments were next performed to determine if the newly identified targets of BldD, *prsI* and *arsI*, were spatially regulated as had been demonstrated for *sigH* (Kelemen *et al.*, 2001). To examine this possibility, attempts were made

Figure 3.33

Fluorescence microscopy of the wild type *S. coelicolor* strain M145 expressing the *bldD-egfp* transcriptional fusion. This work was performed by K. Gislason under the direction of C. Galibois. Shown are *S. coelicolor* M145 carrying pAU257 (pIJ8660 with the *bldD* promoter-*egfp* fusion) or the pIJ8660 vector control. The detector was adjusted to eliminate detection of background fluorescence, so all green fluorescence seen should result from expression of the *egfp* gene under the control of the *bldD* promoter. Strains were inoculated on R2YE at the base of a glass coverslip and grown for the number of days indicated. Panels A and B were grown for two days, C and D were grown for three days, and E and F were grown for four days. Panel G is representative of all images collected from examination of the strain carrying unmodified pIJ8660 (negative control). The order of images is differential interference contrast, fluorescence, and an overlay of the two, from left to right. Aerial hyphae and sporulating structures are indicated by arrows.


2 days
A
B
3 days
D
4 days
E
F
negative control
G

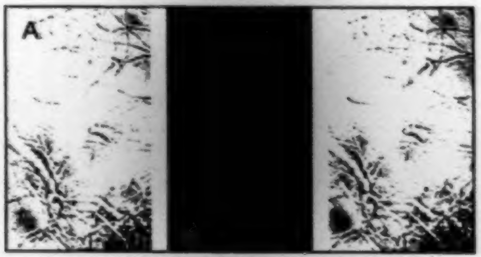
to construct *arsI* and *prsI* promoter-*egfp* transcriptional fusions. For the *prsI-egfp* construct, the procedure was essentially the same as with the *bldD* promoter. Again, C. Fung did the preliminary cloning and K. Gislason finished the cloning and performed the microscopy. A 336 bp fragment extending from 301 bp upstream of the *prsI* start codon to 15 bp inside the coding region (CGA20-CFU1) was amplified, cloned into pCR®2.1 TOPO, generating pAU254, and sequenced. This fragment was transferred as a *Bam*HI - *Xba*I fragment into pIJ8660 (Sun *et al.*, 1999) and the resulting construct was named pAU258. Again, the construct was passaged through *E. coli* strain ET12567/pUZ8002 and conjugated into *S. coelicolor* M145. Chromosomal DNA was isolated from an exconjugant and insertion was confirmed with PCR using CGA20 and KTA-GFP-Rev. The same control strain was used as before (M145 with pIJ8660).

Spores were used to inoculate R2YE solid medium at the intersection with sterilized glass cover slips (Gold Seal®) inserted into the medium at a 45° angle for examination by microscopy, and cultures allowed to grow for two, three and four days. As before, expression was examined through fluorescence and differential interference contrast images, and overlays of the two. As can be seen in Figure 3.34, *prsI* expression was spatially regulated at all three time points examined, being restricted to aerial hyphae and spores. Given that BldD represses *prsI* expression in early vegetative growth, this result was not entirely unexpected, however since the *prsI* transcript is expressed before the erection of aerial hyphae, at least some expression was expected in the substrate mycelia. Perhaps the failure to observe this expression reflects the sensitivity of

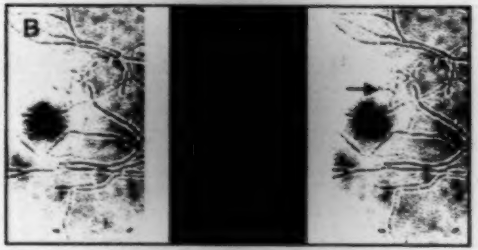
Figure 3.34

Fluorescence microscopy of the wild type *S. coelicolor* strain M145 expressing the *prsI-egfp* transcriptional fusion. This work was performed by K. Gislason under the direction of C. Galibois. Shown are *S. coelicolor* M145 carrying pAU258 (pIJ8660 with the *prsI* promoter-*egfp* fusion) or the pIJ8660 vector control. The detector was adjusted to eliminate detection of background fluorescence, so all green fluorescence seen should result from expression of the *egfp* gene under the control of the *prsI* promoter. Strains were inoculated on R2YE at the base of a glass coverslip and grown for the number of days indicated. Panel A was grown for two days, B was grown for three days, and C was grown for four days. Panel D is representative of all images collected from examination of the strain carrying unmodified pIJ8660 (negative control). The order of images is differential interference contrast, fluorescence, and an overlay of the two, from left to right. Aerial hyphae and sporulating structures are indicated by arrows.

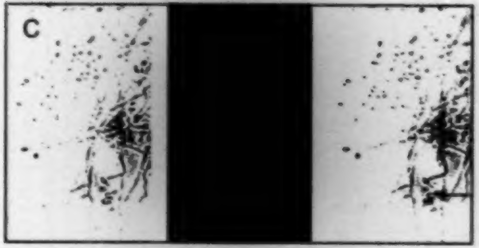
2 days



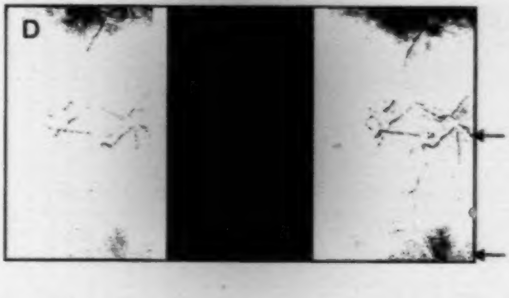
3 days



4 days



negative control



this type of analysis. Regulation of *prsI* transcription will be discussed further in Section 4.1.

3.7.3 Examination of spatial regulation of *arsI* transcription

Spatial regulation of *arsI* was tested in conjunction with that of *prsI*. Again, C. Fung began the work, and K. Gislason continued and finished it. Construct design and testing was done as described above (3.7.1). A 299 bp fragment extending from inside the coding region of the divergently transcribed gene *sigI* to four base pairs inside the *arsI* coding region (CFU3-2) was amplified, cloned into pCR®2.1 TOPO (Invitrogen), generating pAU260, and sequenced. The fragment was transferred as a *Bam*HI - *Xba*I fragment into pIJ8660 (Sun *et al.*, 1999), resulting in pAU261. The construct was transferred to *S. coelicolor* M145 as described above, and spores from the resulting strain were used to inoculate the plates described above. Surprisingly, no fluorescence above background was seen at any time point, despite multiple attempts. *arsI* was known to be expressed at a reasonable level from the S1 nuclease mapping and the primer extension experiments. On the premise that there might be necessary regulatory elements found within the *arsI* coding region, a second construct was tested. The fragment amplified was CFU3-CGA19, a 352 bp fragment extending 57 bp into the *arsI* coding region. Again, the fragment was cloned into pCR®2.1 TOPO (Invitrogen) to generate pAU262, sequenced, and transferred as a *Bam*HI-*Xba*I fragment to pIJ8660 (Sun *et al.*, 1999) to generate pAU263. Examination of *S. coelicolor* M145 with pAU263 by microscopy again showed no fluorescence after any period of growth.

3.8 *sigQ* as a potential activated BldD target

Expression of *sigQ* was very strongly induced as *bldD* expression was induced, as measured by microarray analysis (Fig. 3.8). The *sigQ* gene encodes an extracytoplasmic sigma factor of unknown function. The chromosomal location of *sigQ* is between the *afsQ* operon (divergent) and *sck13.01c/sco4909* (convergent) (Fig. 3.35). Although no connection seems likely between *sigQ* and the convergent gene (a putative ATP-binding protein), it has been proposed that the two component system encoded by the *afsQ* operon directly regulates *sigQ* transcription (M. Hutchings, posted on ScoDB, http://streptomyces.org.uk/). Therefore, *sigQ* is potentially involved in antibiotic production through regulation by the *afsQ* operon, which is believed to regulate antibiotic production in *S. coelicolor*. Overexpression of the *afsQ* operon in *S. coelicolor* causes actinorhodin production in the presence of an *absA* mutation which results in a global block in antibiotic production (Ishizuka *et al.*, 1992) If this is the case, then *sigQ* would provide a long-sought link between *bldD* and antibiotic production; none of the targets discovered to date have been involved in this process.

3.8.1 Effect of *bldD* induction on *sigQ* transcription and sequence analysis of the *sigQ* promoter region

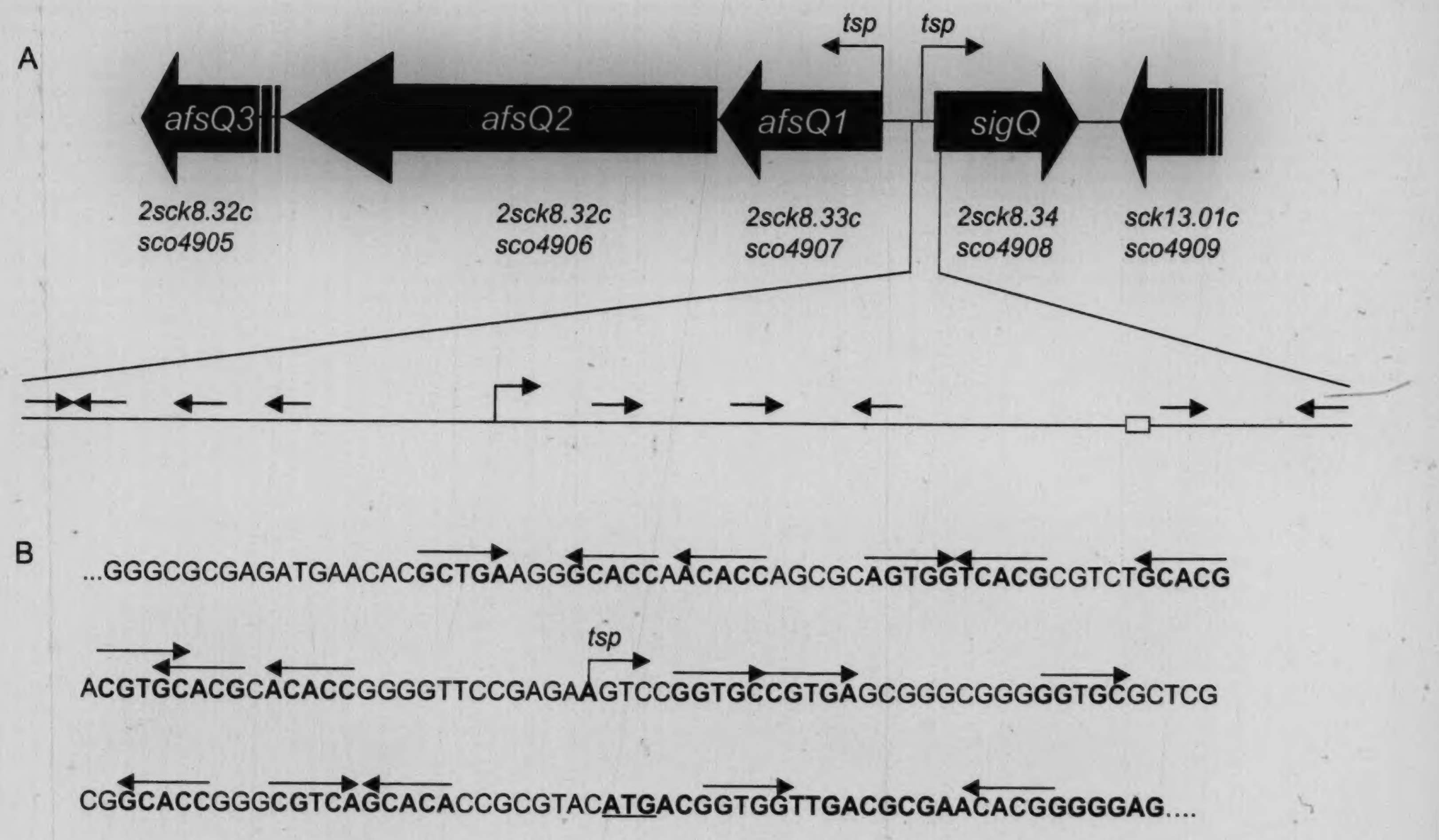
Finding that *sigQ* was a potential target was exciting for more than one reason. Not only was it a potential link between *bldD* and antibiotic production, it also was the first BldD target identified to be induced in the presence of *bldD* induction. Induction of *bldD* resulted in very clear and significant activation of

Figure 3.35

Chromosomal gene environment of *sigQ* and promoter sequence showing putative BldD binding sites.

A) Gene environment of *sigQ*. The *afsQ* operon is shown in teal, *sigQ* is in blue, and the convergent downstream gene is gray. The previously mapped *afsQ* operon transcription start point (tsp) and the *sigQ* tsp are indicated by bent arrows. The expanded section contains the *sigQ* transcription start point, indicated by a bent arrow, and the *sigQ* translation start codon, shown by a box. Putative BldD binding sites are indicated by arrows with the direction of the arrow indicating the direction of the half-site.

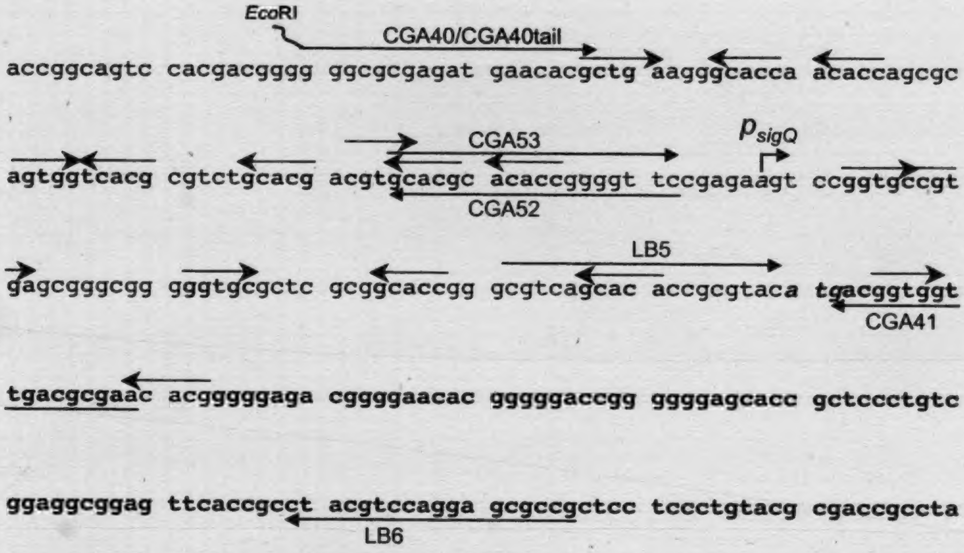
B) Sequence of the area shown in the expanded section of A. Putative BldD binding sites are indicated in red, with the arrow above indicating the direction of the half-site. The coding region of *sigQ* is shown in bold, with the start codon underlined. Again, the *sigQ* transcription start point is indicated by a bent arrow and the corresponding base is indicated in blue.


A
tsp
tsp
afsQ3
afsQ2
afsQ1
sigQ
2sck8.32c
sco4905
2sck8.32c
sco4906
2sck8.33c
sco4907
2sck8.34
sco4908
sck13.01c
sco4909
B
...GGGCGCGAGATGAACACGCTGAAGGGCACCAACACCAGCGCAGTGGTCACGCGTCTGCACG
tsp
ACGTGCACGCACACCGGGGTTCCGAGAAGTCCGGTGCCGTGAGCGGGCGGGGGTGCGCTCG
CGGCACCGGGCGTCAGCACACCGCGTACATGACGGTGGTTGACGCGAACACGGGGGAG....

sigQ in the microarray experiment; the first time point after *bldD* induction showed a six-fold induction of *sigQ* transcription, and the relative transcript levels peaked at 21-fold greater *sigQ* transcript in the induced sample compared to the uninduced sample (Fig. 3.9, E and F). There was no change in the control samples, indicating that this change was a result of *bldD* induction rather than a response to thiostrepton (inducer) addition or other unrelated factors. The initial statistical analysis focused on genes repressed following *bldD* overexpression, as all previously characterized targets of BldD were repressed. Thus, *sigQ* was not identified by the initial analysis. Rather, it was brought to our attention by a personal communication from B. Keijser, who pointed out a possible connection between *ramC* and *sigQ*. At the time, the *bldD* dependence of *ramC* transcription had just been discovered (O'Connor *et al.*, 2002), although sequence analysis indicated that the effect was likely indirect. Examination of the sequence upstream of *sigQ* resulted in the identification of many overlapping putative BldD binding sites located in two clusters. The upstream cluster ranges from 125 bp upstream of the start codon to 80 bp upstream of the start codon, and the downstream cluster extends from 52 bp upstream of the start codon to 24 bp inside the coding region (Figs. 3.35, 3.36). These data together made *sigQ* an object of great interest; it contained more (potential) binding sites than any known BldD target, it was possibly the first induced target of BldD, and it was potentially involved in antibiotic production.

Figure 3.36

Sequence of the *sigQ* coding region and surrounding area. Only the coding strand is shown. Start and stop codons are marked in bold italics, the coding region is bold, and the transcription start point is marked in italic with a bent arrow above. The BldD binding consensus sequences are marked in red, with a red arrow above to indicate the first (forward arrow) or second (reverse arrow) portion of each half site. Primers are marked with an arrow indicating the strand they are derived from; forward primers are marked above the sequence, and reverse primers are indicated below. Wavy lines at the end of some primers indicate engineered restriction sites; the colour of the restriction enzyme is matched to the colour of the primer.



accggcagtc cacgacgggg ggcgcgagat gaacacgctg aagggcacca acaccagcgc

agtggtcacg cgtctgcacg acgtgcacgc acaccggggt tccgaga*a*gt ccggtgccgt

gagcgggcgg gggtgcgctc gcggcaccgg gcgtcagcac accgcgtac***a*** ***tg*acggtggt**

tgacgcgaac acgggggaga cggggaacac gggggaccgg ggggagcacc gctccctgtc

ggaggcggag ttcaccgcct acgtccagga gcgccgctcc tccctgtacg cgaccgccta

ccacctcacc ggcgaccgct tcgaggccga ggacctgctg cagagcgcgc tgttctcgac

ctaccgggcg tgggaccgga tcagtgacaa ggcggcggtc ggcggatacc tccgccgcac

catgaccaac ctgcacatca gcgcctggcg gcgccgcaag ctcaacgagt acccgaccga

ggaactgccg gagacgcccg gcgacatgga cgcgatgcgt ggcaccgagc tgcgcaccgt

cctgtggcag gcgctcgccc ggctgcccga actccagcgc accatgctgg tcctgcgcta

ctacgagggc cgcacggacc cggagatcgc ggacatcctc gacatcagcg tcggcacggt

gaagtccagc atctggcgct cgctgcgccg gctgcgcgag gacgaggtcc tcagcttcgg

ccgtgaccag gaggacgcct tcggcgaact ggtcgcc*tga* aggcacgggc cgggggacc

ggcccgaggg ggcggggctc cgtgaaaaaa gtttcgagcc ggggtgtatc agggggaggg

3.8.2 Determination of the *sigQ* transcription start point

High-resolution S1 nuclease mapping was carried out to determine the transcription start point of *sigQ*. The probe designed was a DNA fragment designated CGA40-tail - CGA41, a 180 bp PCR product which extends from 150 bp upstream of the start codon to 20 bp inside the coding region, with an *Eco*RI site at the 5' end of CGA40-tail to provide the non-homologous upstream sequence. While work was in progress, work in the lab of C. Kao at Stanford University revealed the transcription start point of *sigQ* (N. Karoonuthaisiri, personal communication) and indicated a transcription start point 62 bp upstream of the start codon (Fig 3.36). The single mapping experiment performed with CGA40-tail - CGA41 was of limited success, but there was arguably a band corresponding to the transcription start point identified by N. Karoonuthaisiri (data not shown). Interestingly, the two clusters of putative BldD binding sites found by sequence analysis are located on either side of the transcription start point. One ranges from -70 to -18 and the other from +10 to +86 relative to it. Both clusters of putative BldD binding sites are in a position more typical of repressor binding, as the downstream cluster is positioned well for steric hindrance of the RNA polymerase and the upstream cluster overlaps both the putative -10 and -35 sites. The positions of the two clusters contrasted with the microarray data raise the possibility that BldD may repress *sigQ* at some stages of the life cycle and activate *sigQ* at others, although further analysis is needed to examine this possibility.

3.8.3 Binding of BldD to the *sigQ* promoter region

Given the multiplicity of potential binding sites upstream of the *sigQ* promoter region, an attempt was made to simultaneously prove BldD binding to the region and determine which cluster of binding sites is bound. Initial experiments used a 170 bp probe (CGA40-41) that encompassed all but one of the putative BldD binding sites. EMSAs using this probe showed strong binding of probe by BldD (Fig. 3.37), with a K_D of 0.3 μM, comparable to BldD binding to its own promoter (0.4 μM). The binding was shown to be a specific interaction between BldD and the *sigQ* promoter by addition of a 500-fold excess of unlabelled probe which abolished the lower (first-appearing) BldD-DNA band, although some of the upper band remained. Further experiments demonstrated that BldD binds more than one region within the *sigQ* promoter. One probe, CGA40-52, was designed to encompass the region from 88 bp to five base pairs upstream of the transcription start point (83 bp probe), and the other probe, CGA53-41, 105 bp long, began 28 bp upstream of the tsp and continued to 20 bp inside the coding region. These probes separately encompassed the two clusters of binding sites, although the CGA53-41 does not include the most downstream putative BldD binding site. Probes were end-labelled and incubated with 0.5, 1 or 2 μM BldD, then loaded on a running 8% polyacrylamide gel with glycerol to separate the bound from the unbound probe (Fig. 3.37). The two probes were bound with approximately equal strength; the K_D for CGA40-52 was 1.3 μM while the K_D of BldD binding to CGA53-41 was 1.4 μM, indicating that both regions are necessary for optimal binding of BldD to the *sigQ* promoter.

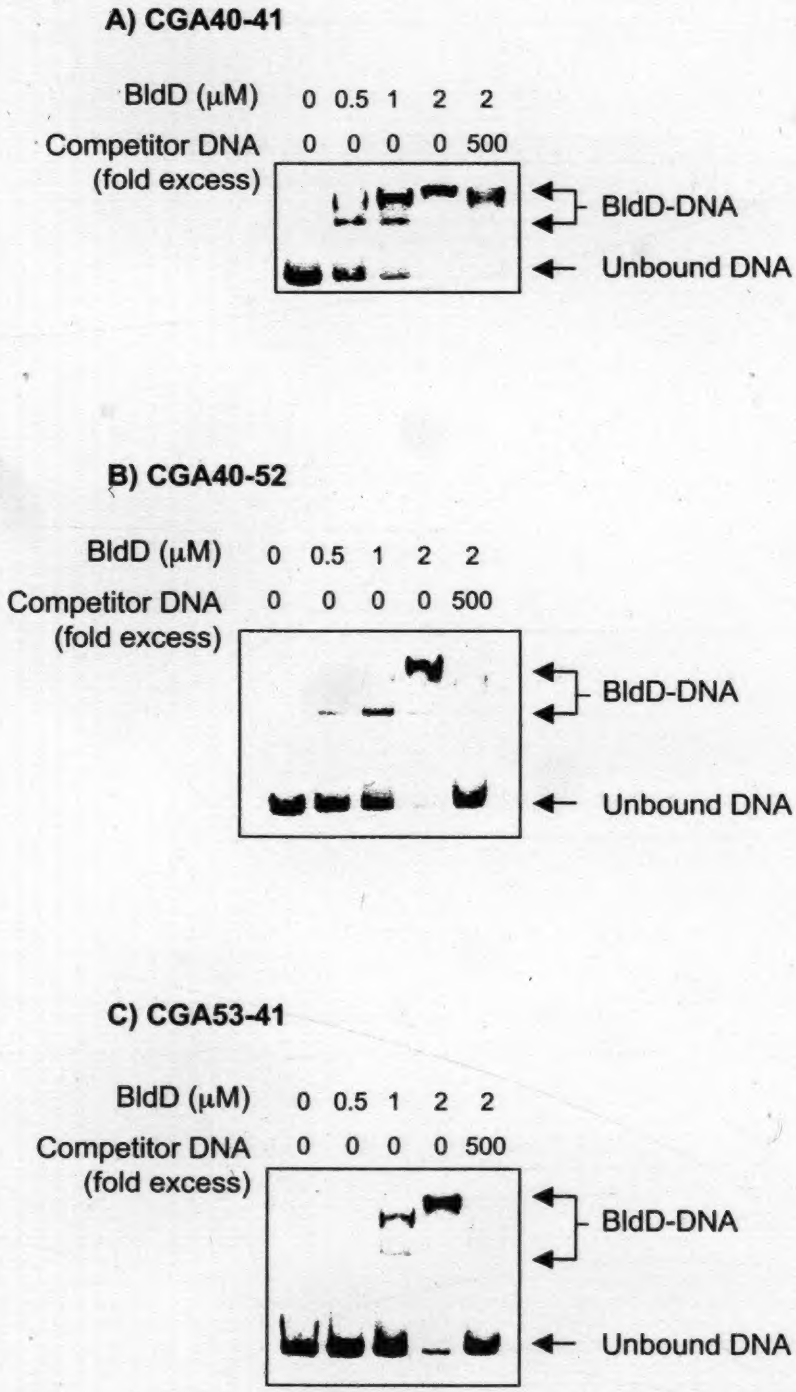
Figure 3.37

BldD binds to the *sigQ* promoter region. EMSAs were performed with three different probes. In all cases, an 8% polyacrylamide gel containing glycerol was used to resolve the bands, and the dried gels were exposed to a phosphor screen to visualize the bands. Probes were end-labelled, then were incubated with increasing amounts of BldD, as indicated. In all cases the final lane contains BldD incubated with a mixture of hot probe and a 500-fold excess of unlabelled probe (competitor DNA). Bound and unbound DNA are indicated by arrows, and the concentration of BldD can be found at the top of the gels.

A) CGA40-41 is a 170 bp probe amplified from cosmid 2SCK8 and spanning the region from 88 bp upstream of the *sigQ* transcription start point to 20 bp inside the coding region.

B) CGA40-52 is 83 bp long and runs from 88 bp upstream of the *sigQ* transcription start point to six base pairs upstream of the transcription start point.

C) CGA53-41 starts 23 bp upstream of the *sigQ* transcription start point and ends 20 bp inside the coding region.


A) CGA40-41
BldD (μM) 0 0.5 1 2 2
Competitor DNA (fold excess) 0 0 0 0 500
BldD-DNA
Unbound DNA
B) CGA40-52
BldD (μM) 0 0.5 1 2 2
Competitor DNA (fold excess) 0 0 0 0 500
BldD-DNA
Unbound DNA
C) CGA53-41
BldD (μM) 0 0.5 1 2 2
Competitor DNA (fold excess) 0 0 0 0 500
BldD-DNA
Unbound DNA

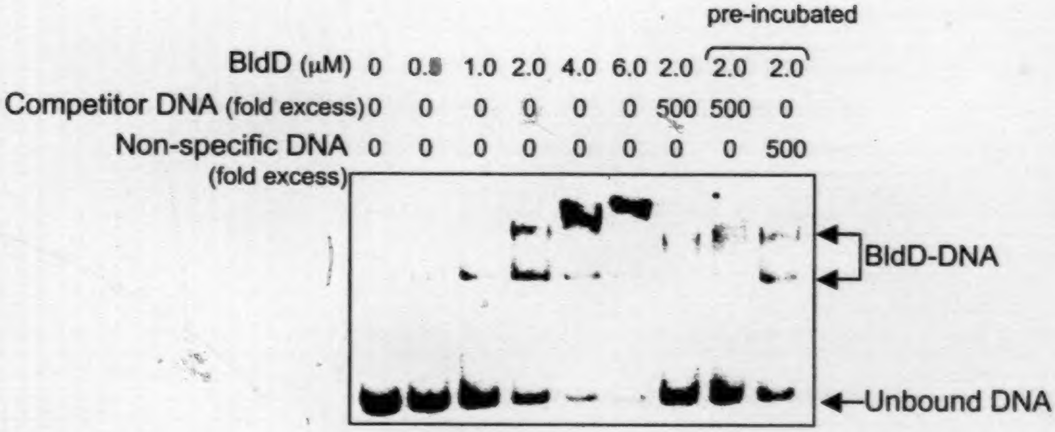
Specificity of binding was confirmed by addition of a 500-fold excess of unlabelled probe, which abolished binding in both cases. Each of the three probes was used in at least two separate experiments. Because of initial problems demonstrating specificity through competition with unlabelled probe, pre-incubation with unlabelled probe was attempted. BldD was pre-incubated at 30°C for half of the binding time (ten minutes) with unlabelled probe before addition of labelled CGA40-52 (Fig. 3.38). Pre-incubation made no difference to the effect of addition specific competitor (reversal or prevention of BldD binding to labelled probe), and so was not used in any other assays. In this same assay, a 500-fold excess of unlabelled DNA from within the *bldD* coding region (BKL41-MAE5) was added to one tube to demonstrate specificity of binding. This fragment was used as a non-specific competitor because BldD was not expected to bind to the *bldD* coding region. No difference in the proportion of DNA bound was seen (Fig. 3.38), given the same amount of BldD, indicating that binding of BldD to the *sigQ* fragment was due to specific interactions and not sequence-independent protein-DNA interactions.

3.8.4 PCR-targeted mutagenesis of *sigQ*

A gene disruption was attempted to help address the role of *sigQ* in *S. coelicolor* antibiotic production and development. This work was done by L. Bui as a summer student under the direction of C. Galibois, using the method of Gust *et al.* (2003), illustrated in Figure 2.2. The apramycin cassette was amplified with the primers CGA54 and CGA55 (Table 2.4), which each contain a 39 bp extension matching sequence flanking the *sigQ* gene to facilitate recombination

Figure 3.38

sigQ EMSA with CGA40-52 (83 bp). An 8% polyacrylamide gel containing glycerol was used to resolve the bands, and the dried gel was exposed to a phosphor screen to visualize the results. The probe was end-labelled, then incubated twenty minutes with increasing amounts of BldD, as indicated in the figure. In the final two lanes, all reagents but the radiolabelled probe were first incubated for ten minutes; addition of the probe was followed by another ten minutes of incubation. Competitor DNA was a 500-fold excess of unlabelled probe, and non-specific DNA was the same excess of a fragment internal to the *bldD* coding region (BKL41-MAE5). Bound and unbound DNA are indicated by arrows, and the concentration of BldD can be found at the top of the gels.


pre-incubated
BldD (μM) 0 0.5 1.0 2.0 4.0 6.0 2.0 2.0 2.0
Competitor DNA (fold excess) 0 0 0 0 0 0 500 500 0
Non-specific DNA (fold excess) 0 0 0 0 0 0 0 0 500
BldD-DNA
Unbound DNA

and replacement of the wild type *sigQ* gene with the apramycin cassette. The product was then introduced by electroporation into *E. coli* BW25113 cells containing the cosmid 2SCK8 (Redenbach *et al.*, 1996) and the inducible λ_{RED} plasmid pIJ790. Integration of the PCR product into the *sigQ* gene on cosmid 2SCK8 was confirmed by digestion with *Eco*RI and *Xho*I (Fig. 3.25). The disrupted cosmid was transformed into the non-methylating *E. coli* strain ET12567 containing the helper plasmid pUZ8002, and then conjugated into the wild type *S. coelicolor* strain M600. Only three exconjugants were obtained and all were found, when grown on fresh plates, to be kanamycin resistant. Two of these were used to isolate unigenomic spores, and the spores were plated to extinction for single colonies. After several generations in the absence of selection, apramycin resistant but kanamycin-sensitive colonies were isolated that had three distinct and stable phenotypes. Several colonies exhibited a wild type phenotype, one isolate produced no red pigment, and one produced no aerial hyphae and no blue pigment. Chromosomal DNA was isolated from representatives of each of these, and from a kanamycin-resistant isolate, and digested with *Bam*HI in preparation for Southern analysis. Three probes were used: the apramycin resistance cassette, the entire cosmid (2SCK8) and a gene-specific probe, in this case the 170 bp PCR-amplified CGA40-41. Southern analysis showed that all of the kanamycin sensitive isolates contained the apramycin resistance cassette, lacked the wild type gene, and had no gross chromosomal rearrangements in the vicinity of *sigQ* (Fig. 3.39).

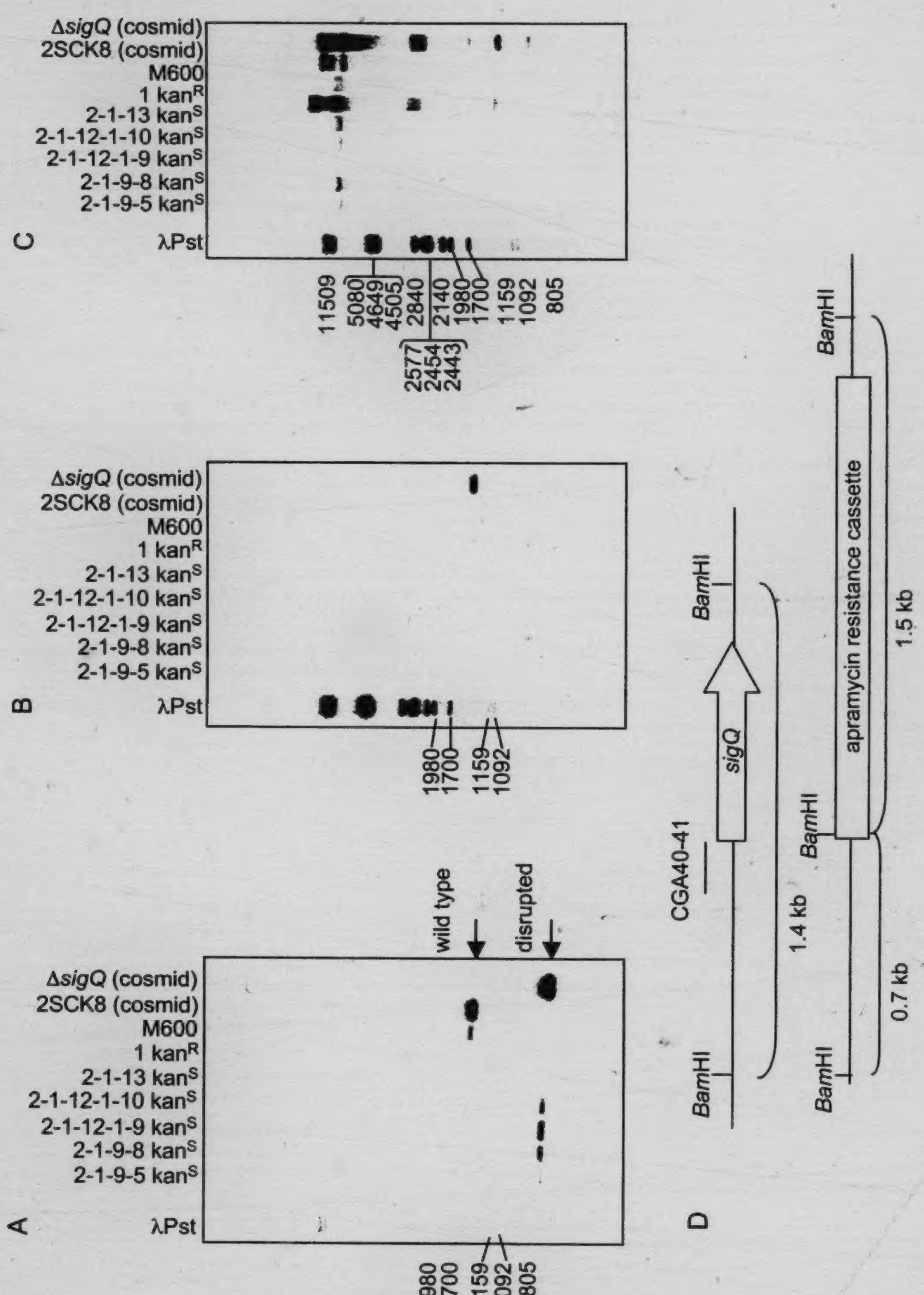
Figure 3.39

Confirmation of replacement of the wild type *sigQ* gene with the apramycin resistance cassette in the chromosome of *S. coelicolor*. This work was performed by L. Bui under the direction of C. Galibois. Chromosomal DNA was isolated from mutants and digested with *Bam*HI before separation by electrophoresis on an agarose gel, transfer to a nylon membrane, and hybridization with a radioactively-labelled probe (see below). The marker used in all cases was lambda phage DNA digested with *Pst*I (λPst).

A) The gene-specific probe used was a 170 bp PCR product (CGA40-41) which covers most of the BldD binding sites upstream of and just inside *sigQ*. Shown are the hybridizing wild-type (larger) and disrupted (smaller) fragments resulting from the restriction digest.

B) The apramycin cassette was used as a probe to confirm the presence of the cassette in the double (kanamycin sensitive) and single (kanamycin resistant) crossover mutants. Similarly digested chromosomal DNA from the wild type *S. coelicolor* (M600) and the undisrupted cosmid (2SCK8) were included as negative controls. The disrupted cosmid (Δ*sigQ*) was included as a positive control.

C) Wild type cosmid (2SCK8) was used as a probe to confirm the absence of any gross chromosomal rearrangements caused by the recombination that resulted in replacement of *sigQ* with the apramycin resistance disruption cassette.


A
B
C
D
ΔsigQ (cosmid)
2SCK8 (cosmid)
M600
1 kanR
2-1-13 kanS
2-1-12-1-10 kanS
2-1-12-1-9 kanS
2-1-9-8 kanS
2-1-9-5 kanS
λPst
wild type
disrupted
11509
5080
4649
4505
2840
2577
2454
2443
2140
1980
1700
1159
1092
805
BamHI
sigQ
CGA40-41
1.4 kb
apramycin resistance cassette
1.5 kb
0.7 kb

In order to determine which of the three was the genuine Δ*sigQ* phenotype, a construct was generated containing a wild type copy of *sigQ* in an attempt to complement the phenotype. First, the cosmid was digested with *Bam*HI and shotgun cloned into pBluescript. The resulting clones were screened by colony hybridization using the 170 bp PCR product amplified by CGA40 and CGA41 (Table 2.4). The resulting plasmid - pBluescript with the 1.4 kb fragment containing the entire *sigQ* coding region - was named pAU250. pAU250 and pSETΩ (O'Connor *et al.*, 2002) were digested with *Bam*HI, and the *sigQ* fragment ligated into pSETΩ to generate pAU251. While this work was in progress, N. Karoonuthaisiri in the lab of C. Kao at Stanford University generated a null mutant of *sigQ* (N. Karoonuthaisiri, personal communication). Her mutants all displayed a wild type phenotype, so no further work was done to characterize the mutants generated by L. Bui. Non-wild type phenotypes can arise from rare spontaneous mutations, and the antibiotic-defective mutants were assumed to contain spontaneous mutations that had appeared during the culturing.

3.9 ChIP analysis of putative BldD targets

ChIP (chromatin immunoprecipitation) is a method for identifying target promoters for transcription factors. It was used in this work to show BldD binding in the wild type *S. coelicolor* strain M600 to known and newly described targets. Experiments were carried out on both solid- and liquid-grown cultures. Cells were harvested after different times (liquid culture) or stages of development (solid culture) in order to see if there were gross differences in BldD binding at different stages of development. All ChIP experiments were carried out in M600 using the M600 derivative, Δ*bldD*1, as a negative control (Elliot *et al.*, 2003b)). Only qualitative information on binding (presence vs. absence) was obtained, given that no attempt was made to strictly quantify the amplification from the immunoprecipitated DNA. An overview of the method can be found in Figure 2.1.

3.9.1 Liquid Culture ChIP

Due to the unexpected results from the liquid-grown cultures (see below for details), each time point was only harvested once. Thus, all results from these experiments can only be considered preliminary.

3.9.1.1 Experimental setup

Initial ChIP (chromatin immunoprecipitation) experiments were carried out based on the literature, using total cell extracts from liquid culture-grown cells (Jakimowicz *et al.*, 2002; Shin and Groisman, 2005; Solomon and Varshavsky, 1985). In the experiments in this work, cultures were grown in TSB/YEME (Section 2.1.4). Antifoam was added to a final concentration of 0.01% to the first set of cultures, as is routinely done in R2YE broth; however, antifoam was found to severely inhibit growth and was not used in subsequent cultures. Cells were washed to remove glycine present in the culture medium, and then any bound proteins were crosslinked to the DNA using formaldehyde. Sonication was then used to rupture the cells and shear the DNA prior to immunoprecipitation with anti-BldD antibody (section 2.6.3). Samples were digested with Proteinase K and extracted with phenol/chloroform after immunoprecipitation, but no other manipulations were performed prior to PCR amplification of any BldD target

DNAs that had been immunoprecipitated by virtue of being crosslinked to bound BldD.

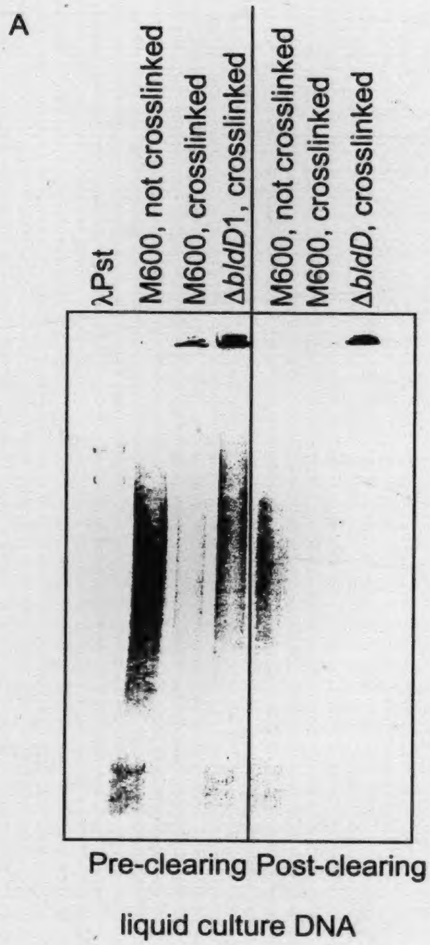
Due to the nature of the procedure, some non-specific precipitation of DNA was inevitable due to co-sedimentation of non-specific DNA with cell debris, or antibody-independent precipitation of very large pieces of unbound DNA. Therefore, measures were taken to minimize background. This involved subjecting total cell extracts to a second round of sonication to further shear the DNA after cell debris had been removed by centrifugation. This met with limited success when viewed after agarose gel electrophoresis; the uncrosslinked DNA was seen to be sheared into mostly 500-1000 bp fragments but a proportion of the crosslinked DNA remained in pieces larger than the largest marker band (11509 bp) with a significant portion remaining in the wells (Fig. 3.40). DNA from the *bldD* deletion strain proved particularly resistant to shearing, resulting in more background amplification than expected in the PCR due to antibody-independent precipitation of large DNA fragments. All sonication was carried out at setting one on the Branson Sonifier 450 (Section 2.6.3) to avoid foaming, however other work (L. Bui, personal communication) found no greater degree of DNA shearing from using a higher setting. Because of this issue, the number of cycles used in PCR amplification was reduced from the standard 30 (end-point PCR) to 25 (semi-quantitative PCR). In this way, it was somewhat clearer which amplified products resulted from DNA precipitated specifically and in high levels, and which products resulted from the trace amounts of DNA present because of insufficient shearing.

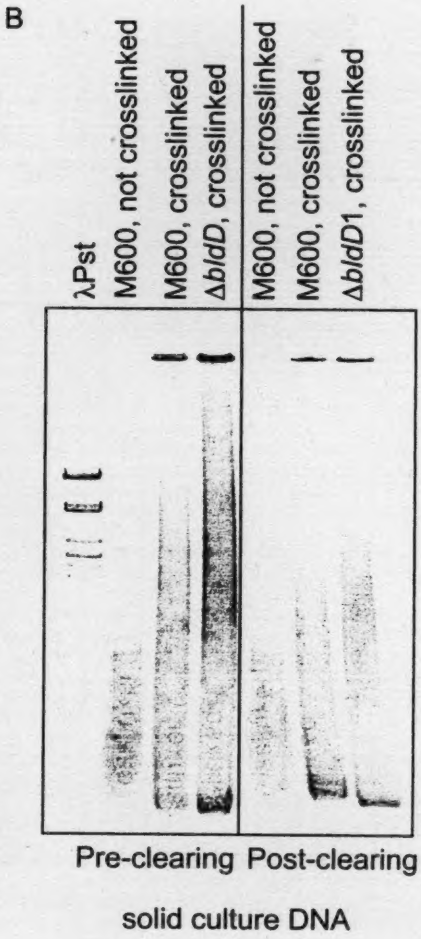
Figure 3.40

Agarose gel electrophoresis of sonicated DNA fragments from total cell-free extracts of *S. coelicolor* generated for ChIP experiments. Samples on the left side of each gel were aliquotted after the initial sonication but prior to clearing by centrifugation to sediment cellular debris (labelled as pre-clearing). Samples on the right side of each gel were cleared by centrifugation then sonicated again (labelled as post-clearing).

A) Typical shearing of DNA isolated from liquid-grown cultures.

B) Typical shearing of DNA isolated from solid-grown cultures.


A
λPst
M600, not crosslinked
M600, crosslinked
Δ*bldD*1, crosslinked
M600, not crosslinked
M600, crosslinked
Δ*bldD*, crosslinked
Pre-clearing
Post-clearing
liquid culture DNA


B
λPst
M600, not crosslinked
M600, crosslinked
Δ*bldD*, crosslinked
M600, not crosslinked
M600, crosslinked
Δ*bldD*1, crosslinked
Pre-clearing
Post-clearing
solid culture DNA

Prior to PCR on ChIP DNA, primer pairs were tested to ensure clean amplification of target genes, with minimal non-specific amplification. Primers were chosen to amplify a region of DNA upstream of the coding region and encompassing the BldD binding region. Target fragment size was 100-300 bp, though primers amplifying a 353 bp region (CGA21-19) were used for *arsI* because extremely small fragments tended to result in an unacceptable level of non-specific amplification. The size range of fragments amplified in the final ChIP analysis was 101 bp to 353 bp.

Each ChIP run generated four samples of DNA (Fig. 2.1). Total DNA was isolated from sonicated, crosslinked cell extracts of M600. No immunoprecipitation was performed on this sample, so it functioned as a positive control for successful amplification of target DNA from crosslinked, sonicated cell extracts. The concentration of target DNA in of this sample was much higher than that of the other samples, so the total DNA was diluted 1/50 before use as template for PCR amplification. A negative control was provided by the extract from Δ*bldD*1 (*bldD* null mutant), grown in parallel with its congenic parent M600. No BldD protein is present in Δ*bldD*1 so any amplification seen in this sample defined the background level of non-specific immunoprecipitation of DNA in the experiment. Another intended negative control was the M600 extract immunoprecipitated without prior crosslinking. The binding of BldD to target DNA in the living cell was not expected to be sufficiently stable to survive sonication and centrifugation, thus a target-specific band above background was not anticipated. The test sample was the M600 extract that was both crosslinked and immunoprecipitated. The crosslinking should reversibly attach BldD to the DNA target, and the immunoprecipitation should result in isolation of only the DNA that is bound by the BldD. All samples were heated to reverse protein-DNA crosslinks before final purification and PCR amplification.

3.9.1.2 Demonstration of *in vivo* binding by BldD

Known BldD targets were used as positive controls in the ChIP experiments. MAE11-12 amplified a 257 bp region extending from 169 bp upstream of the *bldD* transcription start point (tsp) to 27 bp inside the *bldD* coding region, and MAE71-68 amplified a 208 bp fragment extending from 82 bp upstream of the *bldN* tsp to 126 bp downstream of the tsp. As expected, using both primer sets a strongly amplified band was seen in the positive control lane (total M600 DNA, crosslinked but not immunoprecipitated) while no band was amplified from the Δ*bldD*1 sample (Fig. 3.41). Also as expected, a band was amplified from the test sample (crosslinked, immunoprecipitated M600 DNA). Surprisingly, however, a specific product amplified strongly from what was intended as a negative control: the non-crosslinked, immunoprecipitated M600 DNA. This could indicate very strong binding of BldD to both of these promoters, however it was surprising that the binding would survive sonication. The strong binding was seen through amplification in this sample across four time points for both *bldD* and *bldN*. In the experiments shown in Figure 3.41, amplification from the uncrosslinked sample was seen in all genes tested, although inconsistent amplification from the *bldD* mutant strain (Δ*bldD*1) indicated a degree of non-specific precipitation of DNA. For the most part, amplification from the negative

Figure 3.41

Chromatin immunoprecipitation (ChIP) assay of BldD binding to target DNA in cell extracts from liquid cultures. ChIP experiments were carried out on DNA isolated from liquid [3:2 Super-YEME (Hopwood *et al.*, 1985): Trypticase Soy Broth (Difco)] cultures at four time points (18, 24, 30 and 48 hours of growth). The first three samples for each primer set were generated from DNA isolated from the wild type strain M600 and respectively immunoprecipitated without crosslinking (M600, no XL, IP), crosslinked then immunoprecipitated (M600, XL, IP), and crosslinked without subsequent immunoprecipitation (M600, XL, no IP). The third sample (M600, XL, no IP) was used as the total DNA positive control and diluted 1/50 before PCR. The fourth sample (Δ*bldD*1, XL, IP) served as a negative control and was isolated from the *bldD* deletion strain, Δ*bldD*1, before crosslinking and immunoprecipitation. 25 cycles were used for the PCR amplification.

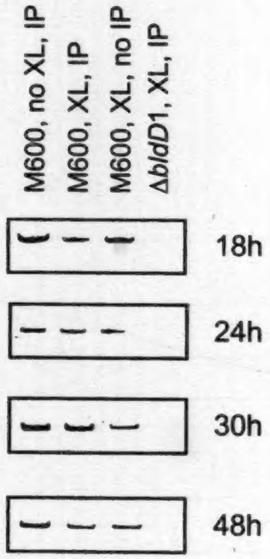
A) *bldD* was amplified with the primers MAE11 and MAE12 (257 bp product).

B) *bldN* was amplified with the primers MAE71 and MAE68 (208 bp product).

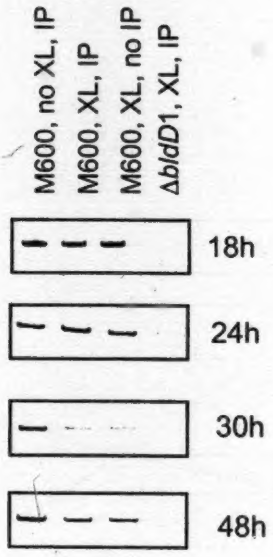
C) *sigQ* was amplified with the primers CGA40 and CGA41 (170 bp product).

D) *hrdB* was amplified with the primers hrdB-F and hrdB-R (101 bp product). *hrdB* is not a BldD target and was included as a negative control.

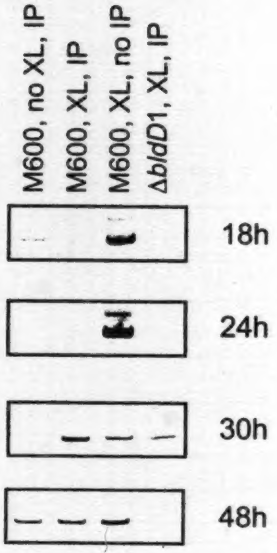
A) *bldD* (MAE11-12)


M600, no XL, IP
M600, XL, IP
M600, XL, no IP
Δ*bldD1*, XL, IP
18h
24h
30h
48h

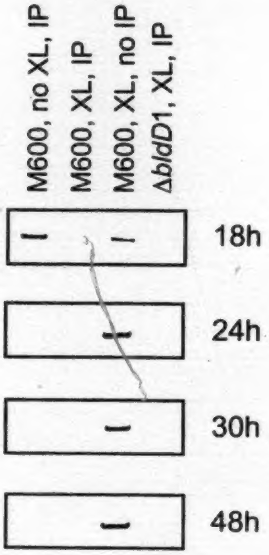
B) *bldN* (MAE71-68)


M600, no XL, IP
M600, XL, IP
M600, XL, no IP
Δ*bldD1*, XL, IP
18h
24h
30h
48h

C) *sigQ* (CGA40-41)


M600, no XL, IP
M600, XL, IP
M600, XL, no IP
Δ*bldD1*, XL, IP
18h
24h
30h
48h

D) *hrdB* (hrdB-F - hrdB-R)


M600, no XL, IP
M600, XL, IP
M600, XL, no IP
Δ*bldD1*, XL, IP
18h
24h
30h
48h

control gene, *hrdB*, was seen consistently only in the total DNA lane and inconsistently observed amplification from other input DNAs will be discussed below.

In the liquid culture ChIP experiments, *hrdB* was used as a negative control, with the primers hrdB-F and hrdB-R used to amplify a 101 bp fragment previously used for S1 nuclease mapping. Although, as expected, amplification was seen only in the total DNA at 24 and 30 hours, there was still significant amplification from the 18 h non-crosslinked, immunoprecipitated M600 DNA sample, and the 48 hour sample had a band in the crosslinked, immunoprecipitated lane (Fig. 3.41). This result was unexpected, but at the time these experiments were done (well before the enzymatic shearing experiments of L. Bui) it was concluded that there was something odd about the samples isolated from liquid culture. This conclusion came from both the *hrdB* results and the unexpected amplification from the non-crosslinked, immunoprecipitated M600 DNA in the *bldD* and *bldN* sets. Therefore further experiments were performed on total cell extracts isolated from cultures grown on solid medium.

3.9.2 Solid Culture ChIP

All previous *in vivo* transcriptional characterization of BldD target gene expression had been performed on RNA isolated from cells grown on solid medium. Therefore, to allow correlations between the results of the ChIP experiments and data from other *in vivo* experiments, all subsequent ChIP experiments were performed on solid-grown cultures.

3.9.2.1 Experimental setup

Much of the experimental setup was the same as for the liquid culture experiments, except that cells were not washed prior to crosslinking, but scraped off of the plates directly into Universal bottles and the formaldehyde mixture used for crosslinking was added immediately. The same issues were found with sonication as with the liquid culture; the crosslinked samples were again highly resistant to shearing (Fig. 3.40). As with the liquid culture experiments, the results are from cell-free extracts harvested only once per time point.

Samples were taken at three time points representing different stages in development of the wild type (M600) strain. At 18 h, aerial hyphae production and antibiotic production had not yet begun; from transcript analysis, this is when BldD exerts its effect by repressing target genes. At 24 hours, aerial hyphae had begun in areas of denser growth, coinciding with red pigment production in those same areas. Cultures grown for 48 hours were mostly sporulated, completely saturated with red pigment, and approximately two thirds covered with blue pigment.

3.9.2.2 Demonstration of *in vivo* binding to known BldD targets

As with the liquid culture ChIP experiments, *bldD* and *bldN* were used as positive controls. These genes were amplified with the primer sets MAE11-12 and MAE71-68, respectively (Fig. 3.42), the same PCR primers used for the liquid culture PCR reactions. Amplification patterns were as expected; there were bands in the positive control lane (1/50 diluted total DNA - crosslinked but not immunoprecipitated) and in the test lane (crosslinked, immunoprecipitated

Figure 3.42

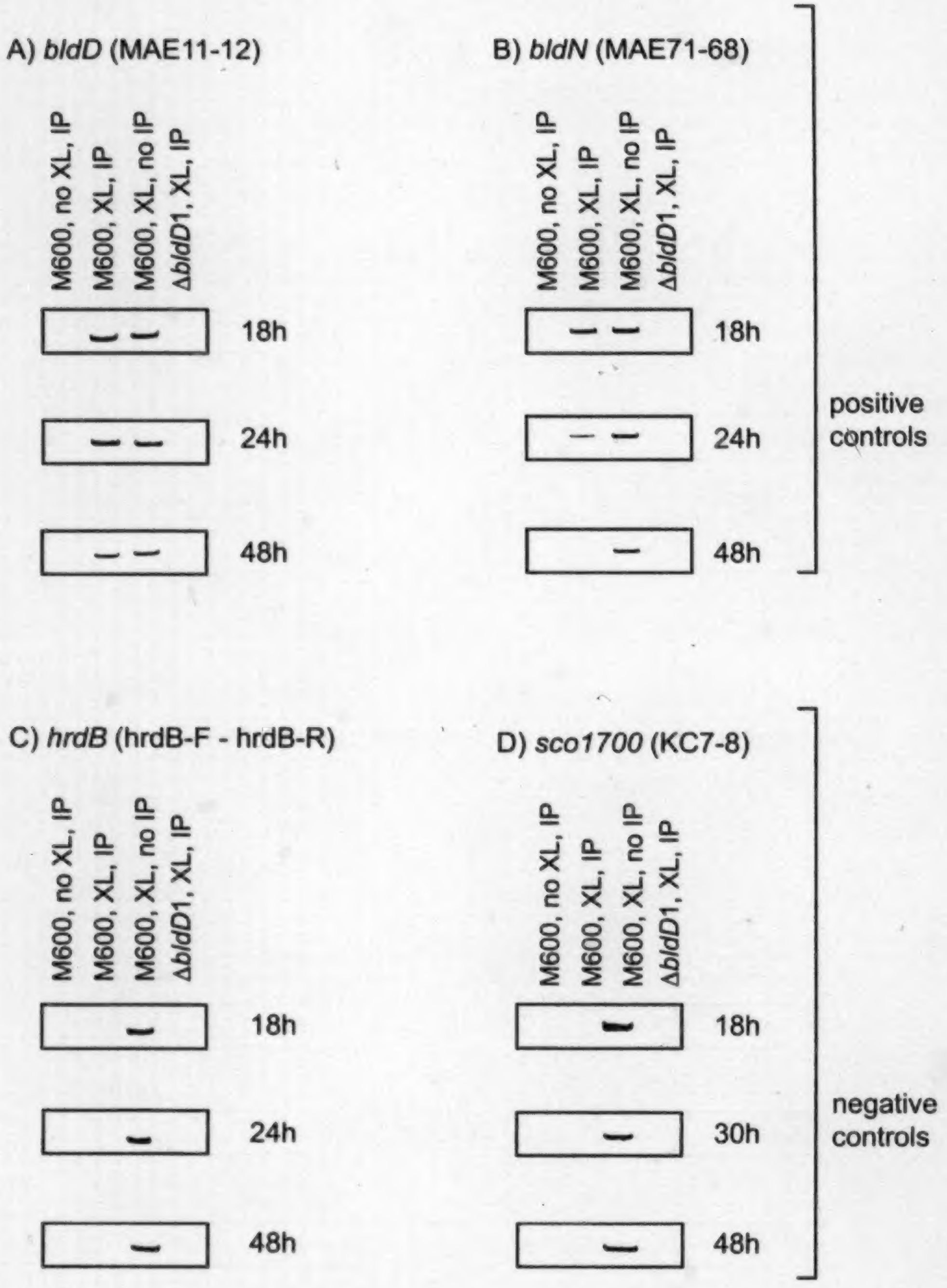
Chromatin immunoprecipitation (ChIP) assay of BldD binding to target DNA in cell extracts from solid cultures. ChIP experiments were carried out on DNA isolated from solid (R2YE) cultures at three time points (18, 24 and 48 hours of growth). The first three samples for each primer set were generated from DNA isolated from the wild type strain M600 and respectively immunoprecipitated without crosslinking (M600, no XL, IP), crosslinked then immunoprecipitated (M600, XL, IP), and crosslinked without subsequent immunoprecipitation (M600, XL, no IP). The third sample (M600, XL, no IP) was used as the total DNA positive control and diluted 1/50 before PCR. The fourth sample (Δ*bldD*1, XL, IP) served as a negative control and was isolated from the *bldD* deletion strain, Δ*bldD*1, before crosslinking and immunoprecipitation. Twenty five cycles were used for the PCR amplification. *bldD* and *bldN* are known BldD targets and were included as positive controls. *hrdB* and *sco1700* are not BldD targets and were included as negative controls.

A) *bldD* was amplified with the primers MAE11 and MAE12 (257 bp product).

B) *bldN* was amplified with the primers MAE71 and MAE68 (208 bp product).

C) *hrdB* was amplified with the primers hrdB-F and hrdB-R (101 bp product).

D) *sco1700* was amplified with the primers KC7 and KC8 (258 bp product).


A) *bldD* (MAE11-12)
B) *bldN* (MAE71-68)
M600, no XL, IP
M600, XL, IP
M600, XL, no IP
Δ*bldD*1, XL, IP
18h
24h
48h
positive controls
C) *hrdB* (hrdB-F - hrdB-R)
D) *sco1700* (KC7-8)
30h
negative controls

DNA) and no, or very faintly observable, bands in the negative controls (uncrosslinked, immunoprecipitated wild type DNA and crosslinked, immunoprecipitated Δ*bldD*1 DNA). This indicated that BldD in the living cell was binding to its own promoter and to that of *bldN*, supporting the *in vitro* evidence of direct binding and *in vivo* evidence of transcriptional dependence. The absence of bands in the negative controls indicated that the binding seen was specific and was not the result of insufficient shearing of template DNA or non-specific DNA precipitation. Results, surprisingly, showed consistent binding to promoters across all time points. BldD is known to repress its own transcription and that of *bldN* only at early time points, so evidence of binding was not anticipated in later development. It is possible that growth even within a single plate was not synchronized such that even at later time points there was some BldD binding to target promoters repressed in earlier development. As well, as stated at the beginning of the section, the data obtained are not truly quantitative; no attempt was made to strictly quantify the amplification from the immunoprecipitated DNA. Small errors in addition (while setting up the PCR reactions or during earlier steps) could result in significant differences in amplification of the final product.

3.9.2.3 Examination of *in vivo* BldD binding to promoter regions of negative control genes

hrdB was again used as a negative control (hrdB-F - hrdB-R). The results in the liquid culture ChIP experiments had unexpectedly shown amplification of *hrdB* in some samples, but this was attributed to the same unknown factors that had resulted in extra bands in the *bldD* and *bldN* reactions. However, when faint bands were seen in the test lane (crosslinked, immunoprecipitated M600 DNA) for all time points isolated from solid culture, it was concluded that *hrdB* was in fact a very poor negative control (Fig. 3.42). It is highly unlikely that BldD regulates the expression of the principal vegetative sigma factor of *S. coelicolor.* The more likely reason that *hrdB* was amplified from the crosslinked immunoprecipitated DNA samples was that this technique is susceptible to precipitation of large complexes through the centrifugation steps (as well as by specific anti-serum:protein interactions) and this might lead to non-specific precipitation of RNA polymerase bound to the very actively transcribed *hrdB* promoter in samples in which both crosslinker and antibody are added. This is supported by the visualization of an amplified, albeit weaker, band in the Δ*bldD*1 lanes as well.

Another gene was then chosen as the negative control. *sco1700* (*sci30a.21c*) is a target of BldG (K. Colvin, personal communication) and has no BldD binding consensus sequence. As well, microarray analysis showed no significant change in its expression following *bldD* induction (Fig. 3.8). The primers KC7-8 were used to amplify a 258 bp fragment extending from -134 to +114 relative to the transcription start point/translation start point of this leaderless transcript (K. Colvin, personal communication). As can be seen in Figure 3.25, unlike *hrdB*, sco1700 behaves as a true negative control. Amplification was seen only in the 1/50 diluted total DNA (crosslinked, but not immunoprecipitated, M600 DNA) and not in any of the other lanes; the same result was seen at all time points. This confirms that BldD does not bind the promoter of *sco1700* and that it makes a good negative control for these types of experiments. When known or putative BldD target genes are subjected to the same experimental conditions, evidence of binding, as seen by amplification from wild type (M600) immunoprecipitated DNA, should indicate actual *in vivo* binding to target promoters.

3.9.2.4 Examination of *in vivo* binding of BldD to *sigQ*

For *sigQ* analysis, a different primer set was used than described above for ChIP analyses with liquid culture extracts. This primer set (CGA40-LB6) amplified a longer fragment of 257 bp that spanned from 88 bp upstream of the *sigQ* transcription start point to 107 bp inside the *sigQ* coding region. Although only a very weak amplified band was visualized in the 24 h cross-linked immunoprecipitated sample, this primer set produced a discernable band in both the 18 h and the 48 h samples without any amplification in the negative control lanes (Fig. 3.43). The apparently weaker binding of BldD to the *sigQ* promoter compared to the *bldD* and *bldN* promoters was unexpected, given the number of putative BldD binding sites upstream of *sigQ*, and given that the K_D of BldD binding to the *sigQ* promoter is 0.3 μM, comparable to *bldD* (0.4 μM). However, lower amplification could result from lower occupancy of the promoter at the stages of development tested, rather than weaker binding. Convincing evidence of BldD binding to *sigQ in vivo* was also found in the work of L. Bui, who saw weak amplification of *sigQ* from immunoprecipitated M600 samples without any amplification from Δ*bldD*1 (Bui, 2006).

Figure 3.43

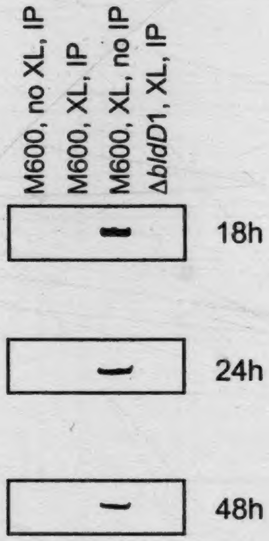
Chromatin immunoprecipitation (ChIP) assay of BldD binding to target DNA in cell extracts from solid cultures. ChIP experiments were carried out on DNA isolated from solid (R2YE) cultures at three time points (18, 24 and 48 hours of growth). Primers used for each gene were CGA43-42 (*prsI*, 162 bp), CGA21-19 (*arsI*, 353 bp) and CGA40-LB6 (*sigQ*, 257 bp). The first three samples for each primer set were generated from DNA isolated from the wild type strain M600 and respectively immunoprecipitated without crosslinking (M600, no XL, IP), crosslinked then immunoprecipitated (M600, XL, IP), and crosslinked without subsequent immunoprecipitation (M600, XL, no IP). The third sample (M600, XL, no IP) was used as the total DNA positive control and diluted 1/50 before PCR. The fourth sample (Δ*bldD*1, XL, IP) served as a negative control and was isolated from the *bldD* deletion strain, Δ*bldD*1, before crosslinking and immunoprecipitation. Twenty five cycles were used for the PCR amplification.

A) *sigQ* was amplified with the primers CGA40 and LB6 (257 bp product).

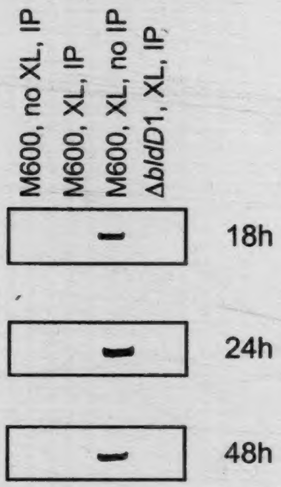
B) *prsI* was amplified with the primers CGA43 and CGA42 (162 bp product).

C) *arsI* was amplified with the primers CGA21 and CGA19 (353 bp product).

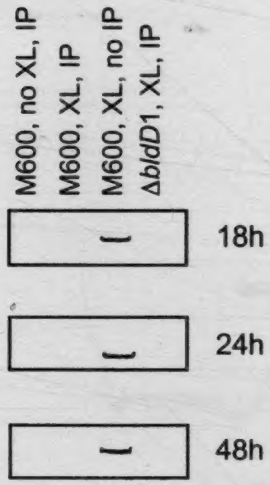
A) *sigQ* (CGA40-LB6)


M600, no XL, IP
M600, XL, IP
M600, XL, no IP
Δ*bldD*1, XL, IP
18h
24h
48h

B) *prsI* (CGA43-42)


M600, no XL, IP
M600, XL, IP
M600, XL, no IP
Δ*bldD*1, XL, IP
18h
24h
48h

C) *arsI* (CGA21-19)


M600, no XL, IP
M600, XL, IP
M600, XL, no IP
Δ*bldD*1, XL, IP
18h
24h
48h

3.9.2.5 Examination of *in vivo* BldD binding to *prsI*

Since the ChIP analysis of solid culture extracts had demonstrated *in vivo* BldD binding to the known targets *bldD* and *bldN*, and to the newly identified target *sigQ*, attempts were also made to use the same conditions to address BldD binding to the *prsI* and *arsI* promoters. Given that BldD bound weakly to *prsI* and *arsI*, weak *in vivo* binding was expected for both *arsI* and *prsI*. As expected, BldD showed very weak binding to the *prsI* promoter region (Fig. 3.43). The primers used (CGA43-42) amplified a 162 bp product found -4 relative to *prsIp2* to 87 bp inside the *prsI* coding region (Fig. 3.11) and were the same primers used to amplify probes for both EMSA and footprinting. Only faint amplification was seen in the test (M600 crosslinked, immunoprecipitated DNA) lane (Fig. 3.43). BldD binding to the *prsI* promoter was seen at all time points, although there was a significant amount of background amplification at 24 hours of growth. Amplification from the sample harvested from the Δ*bldD*1 null mutant indicates that the DNA in other samples was being precipitated because of insufficient shearing of the chromosomal DNA, as well as from specific binding. Taken by itself, this experiment was not strong evidence to prove that *prsI* is a direct, *in vivo* target, but later results from L. Bui in which enzymatic shearing of the DNA was used showed binding of BldD to *prsI* in the absence of any amplification from Δ*bldD*1 (Bui, 2006).

3.9.2.6 Examination of *in vivo* binding of BldD to *arsI*

The primers used to amplify the *arsI* promoter generated the largest fragment amplified in the ChIP experiments. CGA21-19 amplified a 353 bp fragment that extends from within the *sigI* coding region (291 bp upstream of the *arsI* start codon) to 57 bp inside the *arsI* coding region. Weak but visible amplification was seen in all three samples (Fig. 3.43). Unlike *prsI*, no significant background amplification was seen. BldD binding to the *arsI* promoter has been seen *in vitro*, so it is possible that the weak binding observed was a function of the length of the probe; the more successfully used *bldD* and *bldN* probes were 169 bp and 208 bp, respectively. A more likely explanation is that BldD binds more weakly to *arsI* than to its own promoter; BldD binds to its own promoter with a K_D of 0.4 μM (section 3.3.3) but only binds the *arsI* promoter with a K_D of 1.2 μM.

Chapter 4:

Discussion

Chapter 4. Discussion

At the time this work was begun, several targets of BldD regulation were known - *bdtA*, *bldN*, *sigH*, *whiG* and *bldD* itself (Elliot and Leskiw, 1999; Elliot *et al.*, 2001; Kelemen *et al.*, 2001). All but one of these had been found through fairly specific means; for example, the *bld* and *whi* genes (Elliot *et al.*, 2001) were logical targets to screen for potential BldD binding sites, given the evidence of cross-regulation between *bld* genes (Bibb *et al.*, 2000) and the similar role in development. However, once the obvious potential targets had been exhausted, a technique was needed to screen the entire genome of *S. coelicolor*, especially given that the genes known to regulate antibiotic regulation had been examined and none were found to be potential BldD targets (Elliot *et al.*, 2001), and none of the known BldD targets are involved in antibiotic production.

The first global screen to be used was a modified SELEX technique (Elliot, 2000; Elliot *et al.*, 2001). Although this resulted in identification of one new target, *bdtA*, the previously known targets, *bldN*, *sigH* and *whiG*, were never cloned from the selected pool of DNA, indicating that other targets would also be missed by this procedure (Elliot *et al.*, 2001). Other large-scale screening techniques were therefore needed and although two were tested, only one such global screen proved fruitful in this study. The first global screen used was a bioinformatics approach, but it was found to be impractical due to the large variability within the BldD binding consensus. A microarray-based approach was then tried and resulted in the identification of at least three new BldD targets.

4.1 *prsI/arsI* - New Targets found by microarray analysis

prsI (*sce25.07c, sco3066,* putative anti-sigma factor) was among the first targets identified by analysis of the second round of microarray results (*bldD* induction in a *bldD* mutant strain; sections 3.2 and 3.3). Clear repression of *prsI* was seen as a thiostrepton regulated copy of *bldD* was induced, and addition of thiostrepton in the absence of inducible *bldD* had no effect (Figs. 3.8, 3.9 C and D). There was a strong match to the BldD binding consensus upstream of the start codon; three iterations of the first half-site were seen. Transcript mapping indicated two promoters, one seven base pairs upstream of the putative binding site and one a further 42 bp upstream, the latter less strongly expressed. Further examination of the sequence found a weaker match to the BldD binding consensus upstream of the second transcription start site. Binding assays showed that, when tested individually, the downstream site was strongly bound by BldD while the upstream site was bound very weakly. The affinity of BldD binding to a larger fragment containing both sites was approximately the same as the affinity of BldD binding to the downstream site alone. The weak binding of BldD to the upstream site does not indicate an absence of binding *in vivo*; a recent publication (Kim *et al.*, 2006) showed very weak binding to proven binding sites (one each of *bldN* and *whiG*) when these sites were tested in isolation. *bldN* and *whiG* both have two BldD binding sites, demonstrated by DNase I footprinting (Elliot *et al.*, 2001). However, when these sites are tested individually by EMSA, one site in each promoter is not bound at all. The authors suggest that perhaps BldD binds co-operatively: first binding to one site and inducing a bend in the DNA, then binding to the other site (Kim *et al.*, 2006). DNase I footprinting would indicate if both sites upstream of *prsI* are bound by BldD in a larger fragment. Although such analyses were unsuccessful in this work, it is hoped that they may prove less recalcitrant in other hands. BldD binding to *prsI in vivo* was demonstrated by ChIP analysis, although these experiments should be repeated using enzymatic DNA shearing as described by Bui (2006) as background amplification was difficult to control using the methods described in this work.

As a target of BldD, *prsI* represents a different case than *whiG* or *bldN*. Both of the latter genes have only one promoter while *prsI* has two. The immediate assumption is that each binding site should regulate one promoter, but the upstream site (near the less-active promoter) is likely only bound after the downstream site has been occupied. This suggests simultaneous regulation of both promoters by BldD, but then why have two promoters? The most likely explanation is differential regulation by conditions unrelated to BldD binding, such as environmental stress. Addressing this question was beyond the scope of this work but could be tested by harvesting RNA from cells exposed to various stresses and examining the transcriptional response from *prsIp1* and *prsIp2*. Both promoters exhibit similar temporal regulation in the wild type, though *prsIp2* (the weaker promoter) is expressed several hours later than *prsIp1*. All temporal regulation of both promoters is eliminated in a *bldD* mutant.

Transcript mapping of *prsI* showed co-transcription with the upstream gene as well as independent transcription. *arsI* (*sce25.08c, sco3067*) encodes a putative anti-anti-sigma factor which is likely to act with PrsI to regulate the divergently-encoded SigI (Fig. 3.15). An excellent BldD binding site was also found upstream of *arsI* (proven by EMSA, ChIP and DNase I footprinting) and regulation of the *arsI/prsI* polycistronic transcript was also found to be *bldD*-dependent in early growth. Temporal regulation of the long transcript was somewhat more strict than that of the *prsI*-specific transcripts, as transcription of the operon only began at the onset of aerial hyphae production while production of the *prsI*-specific transcripts began prior to the formation of aerial mycelia. Given that there is very little temporal separation in the expression of the three transcripts, differential spatial expression was tested.

BldD is known to spatially, as well as temporally, regulate at least one of its targets (Kelemen *et al.*, 2001). Expression of *sigH* is restricted to aerial hyphae in wild type *S. coelicolor* but present in all tissues in a *bldD* mutant grown on minimal medium with mannitol. To determine if BldD also restricts the expression of *arsI* and *prsI* to specific tissues, the integrating plasmid pIJ8660 (Sun *et al.*, 1999), with the promoters of *arsI*, *prsI* or *bldD* cloned in front of the gene encoding enhanced green fluorescent protein (*egfp*), was used to generate derivatives of the wild type strain M145. As expected, *bldD* itself was expressed in all tissues at two and three days' growth (Fig. 3.33). Interestingly, though, there was no fluorescence above background at four days of growth. This result was not entirely unexpected because previous results had shown that *bldD* transcription peaks in early growth and decreases rapidly thereafter in the wild type strain J1501 (Elliot *et al.*, 1998), although no decrease in BldD protein levels was seen in the same period in the same strain (Elliot, 2000). As well, in another wild type strain, M600, BldD levels have been shown to decrease after 48 hours of growth on solid medium (Bui, 2006).

Expression of *prsI* was restricted to aerial hyphae in the time points tested (2, 3 and 4 days) (Fig. 3.34). This was not expected. *prsI* expression cannot always be restricted to the aerial hyphae; gene expression begins before the formation of aerial structures. One possible explanation for this is a developmental switch; upon initiation of aerial hyphae formation, expression of *prsI* is prevented in the substrate and permitted in the aerial hyphae. Expression of the polycistronic transcript from p_{arsI} would be expected to follow the pattern of *prsI* expression; the function of an anti-sigma factor antagonist is to regulate the activity of the anti-sigma factor, so that *arsI* expression would be needed for expression of genes directed by the PrsI-antagonized sigma factor. Spatial separation of the mono- and polycistronic transcripts could help to modulate the relative levels of PrsI and ArsI, ensuring that there are not too many PrsI molecules for the active ArsI molecules to regulate and allowing fine control of target sigma factor expression. Normal regulation of anti-sigma factor/anti-sigma factor antagonist systems is fine-tuned by post-translational regulation through phosphorylation, indicating that spatial regulation of *prsI* transcription may be only one aspect of the regulation of its function.

The putative product of the *prsI* gene has a predicted kinase domain with homology to the kinase domain of RsbW of *B. subtilis*. RsbW is the anti-sigma factor that regulates σ^B, the primary stress response sigma factor of *B. subtilis*

(Kuo *et al.*, 2004). The anti-sigma factor RsbW binds to and prevents the activity of σ^B, and is in turn bound and hindered by the anti-sigma antagonist RsbV. When σ^B activity is not needed, RsbV is kept in its phosphorylated (inactive) state by the kinase activity of RsbW. Under conditions of stress, however, the phosphatase RsbU becomes activated and removes the phosphate group from RsbV, converting it to its active state (Kuo *et al.*, 2004).

A gene encoding a predicted homologue to the phosphatase RsbU has been found in the *S. coelicolor* genome, although it was not characterized in this study. Microarray data for this gene, *sco0751* (*scf81.10c*), demonstrated that it was repressed during *bldD* induction (Figure A.1, A and B), and examination of the sequence upstream of the start codon found a match to the BldD binding consensus sequence, GGTGC (7) CCACC, indicating that *sco0751* is a putative direct BldD target (see Appendix A.1). In this way, it appears that BldD regulates the transcription of three genes involved in the post translational-regulation of sigma factors - *prsI* (anti-sigma factor), *arsI* (anti-sigma factor antagonist) and *sco0751* (anti-sigma factor antagonist phosphatase). These data raise the possibility that the product of *sco0751* is involved in the post-translational regulation of PrsI and ArsI. Supporting this, the predicted product of *arsI* has significant homology to SpoIIAA, the anti-sigma factor antagonist in *B. subtilis* dephosphorylated by SpoIIE (Yudkin and Clarkson, 2005), and the predicted phosphatase domain of *sco0751* shows homology to that of SpoIIE of *B. subtilis* as well as RsbU (determined using the NCBI Conserved Domain Database) (http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml). This speculation, however, does not address the possibility that the product of *sco0751* instead regulates an unrelated anti-sigma factor antagonist. The target protein of SCO0751 would have to be addressed by further experiments.

A series of experiments to further investigate these relationships would provide further insight into the regulation of development in *S. coelicolor*. The first experiment needed would be determination of protein levels of ArsI and PrsI throughout growth. Although gene expression has been demonstrated, transcript levels do not necessarily correlate to protein levels. Western analysis would demonstrate both expression of the proteins and developmental regulation of translation. This experiment requires overexpression and purification of both proteins from *E. coli* and purification of antibodies to them. The purification techniques optimized and antibodies produced for this experiment would then be useful for further experiments. For example, one could determine if ArsI is regulated by phosphorylation, as is the better-characterized putative anti-sigma factor antagonist BldG (Bignell *et al.*, 2003). The predicted interaction between PrsI and ArsI could be examined by affinity chromatography with purified proteins, using a previously described method (Nussbaum-Shochat and Amster-Choder, 1999) as modified by Bignell (2003) or by chemical crosslinking using extracts from *S. coelicolor* cultures (Alper *et al.*, 1994; Bignell, 2003). The latter experiment could also potentially demonstrate interaction between ArsI and SCO0751.

An important experiment would be to determine if *sco0751* is, in fact, a BldD target. Characterization could proceed as for *prsI* and *arsI*: transcriptional mapping of RNA isolated from wild type and *bldD* mutant strains, followed by EMSA and footprinting experiments if transcription of *sco0751* is *bldD*-dependent. ChIP analysis, using enzymatic DNA shearing as described in Bui (2006), would be useful in demonstrating *in vivo* binding of BldD to the promoter of *sco0751*. Quantitative analysis of the ChIP results would seem to be useful for measuring the strength of BldD binding *in vivo*, however there are many steps involved in purification of DNA from ChIP so that even the most careful analysis would not be able to eliminate all sources of error. As well, ChIP analysis cannot account for occupancy at the promoter, which can be affected by other factors such as presence of other transcription factors, and binding strength of the protein in question to the promoter being tested. Finally, deletion of the coding region of *sco0751* using the method of Gust *et al*. (2003) would demonstrate if this gene has a role in *S. coelicolor* development or antibiotic production, and could be carried out in parallel with the transcriptional mapping.

Although deletion of a number of genes (i.e. *arsI*, *sigI*, and *sigQ*) was accomplished in this work, *prsI* proved resistant to deletion, suggesting that the role of *prsI* may be essential for *S. coelicolor* viability. Fortunately, disruption of a gene is not the only means to examine its function. Overexpression can also show the purpose of a gene. Overexpression of *prsI* was expected to yield a phenotype similar to the disruption phenotype of *arsI*: deficient in aerial hyphae formation. Overexpression of the *prsI* gene should result in the production of far more PrsI than is normally present. This would then be expected to overcome the normal post-translational regulation, and prevent activity of the target sigma factor(s) in the same way as absence of the anti-sigma factor antagonist. Attempts to overexpress *prsI* were unsuccessful and are discussed in Appendix C.

The essential nature of *prsI* is not unprecedented. While no other anti-sigma factors in *S. coelicolor* have been found to be essential, a precedent can be found in *Bacillus subtilis*. As stated earlier, *prsI* is a homologue of *spoIIAB* of *B. subtilis*, which is involved in the complex regulation of sporulation in that organism. *spoIIAB* has proven quite resistant to disruption; the deletion mutation *spoIIAB*D*1* was highly unstable and accumulated suppressor mutations that blocked the function of the target sigma factor σ^F, which is encoded in an operon with its anti-sigma factor (*spoIIAB*) and the cognate anti-sigma factor antagonist (*spoIIAA*). When these suppressor mutations were eliminated by backcrossing, the resulting colonies lysed after one to two days of growth (Schmidt *et al.*, 1990). The authors attributed the effect to inappropriate stimulation of σ^F activity; the uncontrolled activity of σ^F during vegetative growth competed with the primary vegetative sigma factor σ^A and prevented expression of essential vegetative genes. There are examples in *S. coelicolor* where uncontrolled sigma factor activity causes deleterious effects, though not cell death. Deletion of the putative anti-sigma factor-encoding gene *rsuA* causes a *bld* phenotype, as does overexpression of *sigU*, which encodes an apparent extracytoplasmic function sigma factor, but deletion of *rsuA* and *sigU* together has no effect on development (Gehring *et al.*, 2001). Null mutants of *rsrA*, which encodes the anti-sigma factor which regulates the disulphide stress response sigma factor σ^R, are defective in sporulation, an effect which the authors speculate is likewise caused by inappropriate activity of the target sigma factor (Paget *et al.*, 2001). In our case, it is possible that deletion of *prsI* does not simply result in deregulation of the sigma factor encoded nearby (*sigI*) as a null mutant of *arsI* exhibits a *bld* phenotype, suggesting that the target sigma factor(s) of ArsI/PrsI regulation are involved in development. As seen in this work and in other work (Lee *et al.*, 2005), deletion of *sigI* causes no visible effect on cell growth or development. If SigI is regulated by PrsI/ArsI, then there is at least one other sigma factor also regulated.

The observation that BldD does not regulate the transcription of *sigI* was not an unexpected result. This finding is consistent with the work of Viollier *et al* (2003), who saw no difference in *sigI* transcription in a variety of *bld* gene mutants when compared with the wild type. As expected, the *sigI* null mutant has no obvious difference in phenotype when compared to the wild-type strain, either on rich medium or when subjected to osmotic shock on minimal medium (Figs. 3.28 and 3.31). SigI is induced by osmotic shock but there are several other proven or putative sigma factors that are also induced by osmotic shock including σ^H, SigJ and SigX (Viollier *et al.*, 2003), not to mention the master regulator σ^B which directs the response to osmotic shock (Lee *et al.*, 2005). This redundancy means that multiple knockouts would likely be necessary to see an effect on viability following osmotic stress. Given the dramatic differences between the phenotypes of the *sigI* null mutant (essentially wild-type) and the *arsI* null mutant (completely bald), as well as the essential nature of *prsI*, it is clear that BldD regulation through PrsI and ArsI is not limited to regulation of SigI.

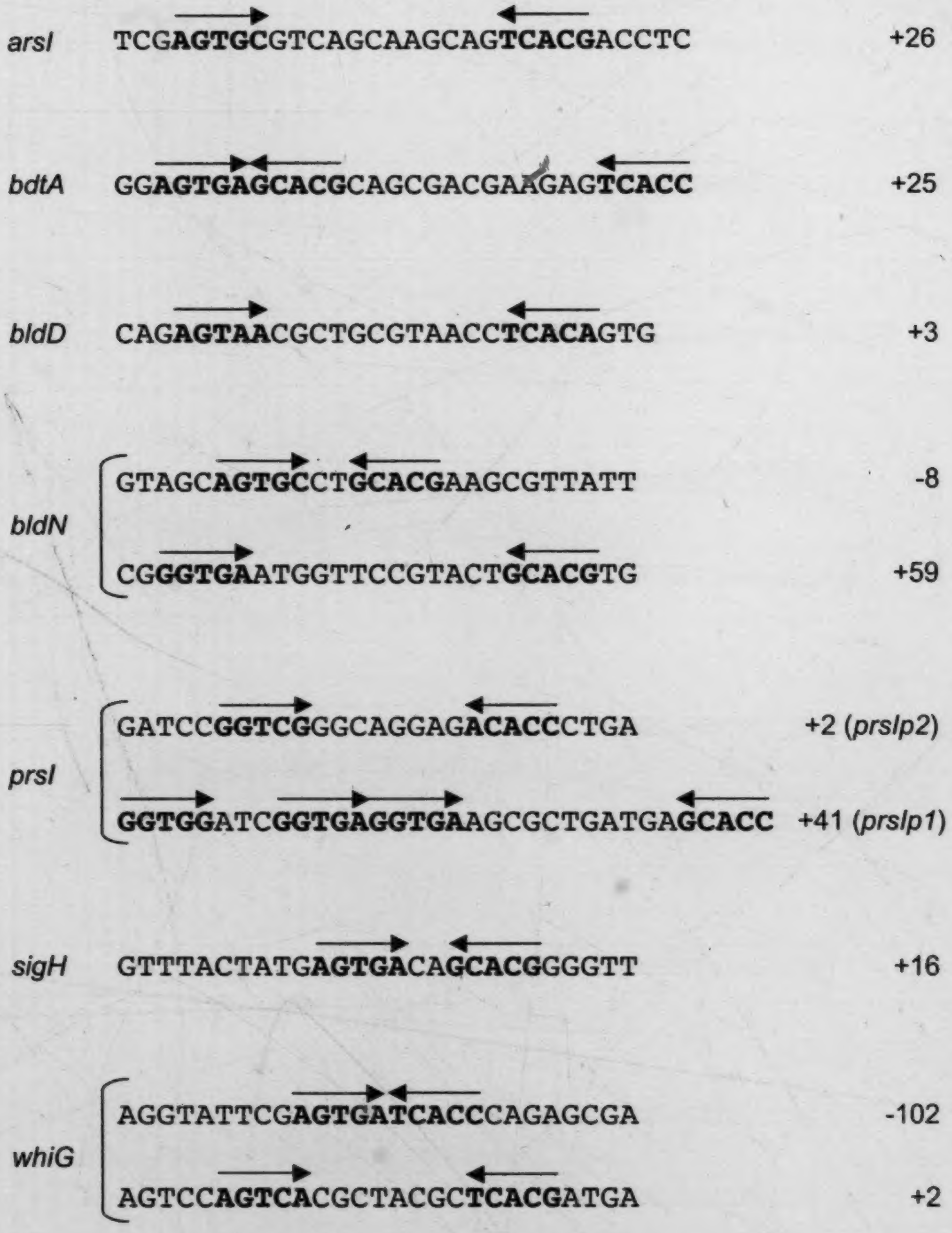
The expression patterns of *prsI* and *arsI*, both temporal and spatial, lend themselves to a number of explanations of their potential role in *S. coelicolor* development. One possibility, although highly speculative, is that PrsI and ArsI regulate sigma factors involved in regulation of PCD (programmed cell death), the orderly dismantling of older mycelia to provide nutrients and structure for developing mycelia (Miguelez *et al.*, 1999). Although most studied in *S. antibioticus*, PCD has also been shown in *S. coelicolor* (Manteca *et al.*, 2006). In *S. antibioticus*, PCD occurs first in the substrate mycelium (during aerial hyphae formation), then later in the basal, nonsporulating aerial hyphae (during sporulation) and is proposed to release nutrients necessary for the next stage of growth while still maintaining the physical structure of the colony (Miguelez *et al.*, 1999). PCD in the substrate hyphae would provide resources necessary for the erection of aerial hyphae, and dismantling of the lower (nonsporulating) aerial hyphae would in turn provide components needed for the full maturation of spores. The *bld* phenotype of *arsI* mutants can be explained through involvement in PCD; without ArsI, expression of genes needed for PCD might be blocked. If true, without PCD there might not be enough nutrients for erection of aerial structures. Premature death within inappropriate tissues, for example when *prsI* is knocked out, could result in early death of the substrate mycelium, preventing any meaningful amount of growth.

This speculation would be supported by EGFP results showing that expression of the polycistronic *arsI/prsI* transcript follows the spatial pattern of *prsI* expression, only hours later. To clarify this statement in the context of this model, expression from p_{arsI} would have to begin at the onset of aerial hyphae expression (as shown), and be restricted initially to the substrate hyphae. At the same time, *prsI* expression would be restricted to the newly formed aerial hyphae (as shown in figure 3.34). In this way, activity of the PCD-specific sigma factor(s) would be possible in the older (substrate) hyphae while prevented in the growing (aerial) hyphae. Unfortunately, *arsI* has proven resistant to EGFP analysis. Experiments with p_{arsI}-driven *egfp* expression, which would provide some support for this proposed model, met with technical difficulties and could not be completed at this time.

A much simpler explanation of the data would appear to be that PrsI and ArsI together regulate a sigma factor which directs expression of genes necessary for erection of aerial hyphae. If true, the resistance of *prsI* to mutation could then be explained by deregulation during vegetative growth of target sigma factor activity, with the mutant phenotype of *arsI* indicating that the target sigma factor(s) is necessary for normal development of the colony. It is odd, though, that *arsI* expression would begin in tandem with aerial hyphae formation; genes involved in formation of aerial structures are more typically expressed before aerial hyphae are visible. Most likely, initiation of *arsI* expression began between the two time points harvested, and so *arsI* was not seen to be expressed prior to aerial hyphae formation. Restriction of *prsI* to the aerial hyphae could indicate that the regulated sigma factor(s) has a role solely in erection of the aerial hyphae and not in their development. Identification of the sigma factor(s) regulated by PrsI and ArsI will help to clarify the role of these proteins in *S. coelicolor* development, as currently the data do not unequivocally support either conclusion.

4.2 Refined BldD binding consensus sequence

As new targets were analyzed, it became possible to refine the consensus sequence for BldD binding to its target promoters. The previously published consensus was AGTgA $(n)_m$ TCACc where n represents any base and the spacer (n) is of variable length (m) (Elliot *et al.*, 2001). Re-examination of the published targets, along with sequence analysis of the *sigH* footprint (Kelemen *et al.*, 2001) and the targets discussed in this work, resulted in a refined consensus of AGTGA (n_m) g/tCACg/c (Fig. 4.1). A graphical representation of the refined consensus (Fig. 4.2) was generated using WebLogo (http://weblogo.berkeley.edu/), a program designed to provide a consensus that describes more effectively the variation in sequence within the sites used to generate a consensus sequence (Crooks *et al.*, 2004). Although the refined consensus appears to be less certain than the previously published consensus, examination of the sequence logos show that conservation at positions three, seven and eight has increased significantly, such that there is absolute conservation in five of the ten bases in the revised BldD consensus binding sequence (Fig. 4.2). Other positions are somewhat less conserved than in the original sequence, but as no attempt was made to weight the revised BldD



arsI TCG**AGTGC**GTCAGCAAGCAG**TCACG**ACCTC +26

bdtA GG**AGTGAGCACG**CAGCGACGAAGAG**TCACC** +25

bldD CAG**AGTAA**CGCTGCGTAACC**TCACA**GTG +3

bldN GTAGC**AGTGC**CT**GCACG**AAGCGTTATT -8

CG**GGTGA**ATGGTTCCGTACT**GCACG**TG +59

prsI GATCC**GGTCG**GGCAGGAG**ACACC**CTGA +2 (*prsIp2*)

GGTGGATC**GGTGAGGTGA**AGCGCTGATGA**GCACC** +41 (*prsIp1*)

sigH GTTTACTATG**AGTGA**CA**GCACG**GGGTT +16

whiG AGGTATTCG**AGTGATCACC**CAGAGCGA -102

AGTCC**AGTCA**CGCTACGC**TCACG**ATGA +2

previously published consensus: AGTgA (n_m) TCACc

refined consensus: AGTGA (n_m) g/tCACg/c

Figure 4.1

BldD binding sites in seven target promoters. *arsI* and *prsI* were first described as BldD targets in this work; the rest have been previously published. (Elliot *et al.*, 2001; Kelemen *et al.*, 2001). All of the binding sites have been confirmed by EMSA; all but *prsI* have also been confirmed by DNase I footprinting. BldD binding sequences are marked in bold, and the orientation of the binding half-site is indicated by the arrow above. Distance of the last base of the sequence from the transcription start point is indicated on the right side of the figure. The refined consensus is from this work.

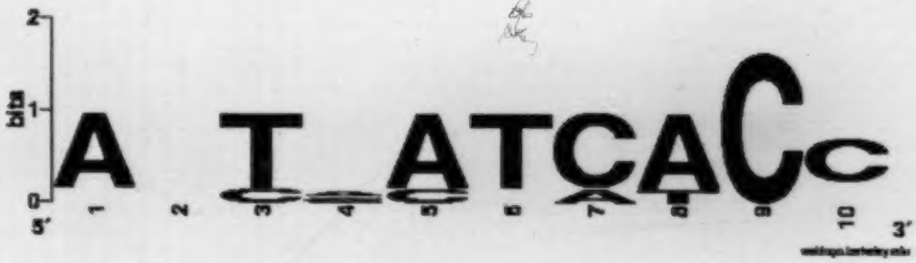
Figure 4.2

Sequence logo illustrating the previously published and refined BldD binding consensus sequences. Sequence logos (generated by WebLogo at http://weblogo.berkeley.edu/) are a graphical representation of alignments of sequences, in this case the BldD binding consensus sequence. The logo is produced from statistical analysis of the level of conservation of the bases at each position within the consensus. The height of the stack at each position represents the overall level of conservation at that position, and the height of each letter within the stack represents the relative frequency of that letter at that position. Intervening sequences between the half-sites were variable in length, and were not included in the analysis. Sequence logos are designed to preserve the information lost by the traditional presentation of a consensus sequence.

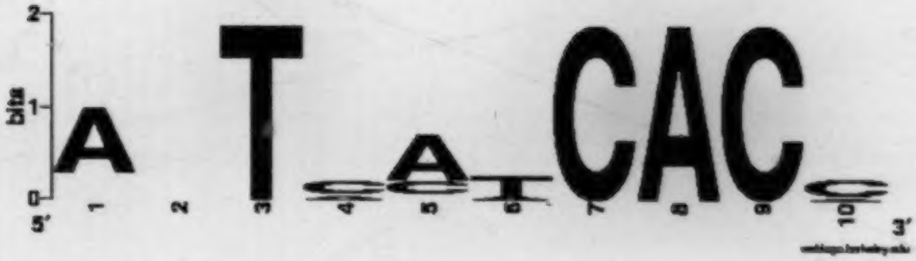
A) Previously published BldD binding consensus.

B) Refined BldD binding consensus.

A


bits
2
1
0
5'
1
2
3
4
5
6
7
8
9
10
3'
weblogo.berkeley.edu

B


bits
2
1
0
5'
1
2
3
4
5
6
7
8
9
10
3'
weblogo.berkeley.edu

binding consensus according to strength of binding, this variation could simply reflect inclusion of less than optimal binding sites. Revised BldD binding sites for known targets are listed in Table 4.1. All revised binding sites, and the site found for *sigH*, are found within the published footprint of BldD binding. The revised BldD binding sequences for *bldD* and *bdtA* had better matches to the better conserved bases in the published consensus, and the sequence chosen for the upstream BldD binding site of *bldN* reflected the absolute conservation of the CAC motif after the re-evaluation of the second half of the binding consensus in the other BldD targets. Although the new consensus is less defined than the old, it reflects the significant variations within known BldD binding sites. No match to the old BldD binding consensus had been found within the region of BldD binding upstream of *sigH* (Kelemen *et al.*, 2001), but there is a perfect match to the refined consensus within the footprinted region (Table 4.1). The binding sites within *sigQ* were not included for two reasons. First, there are so many of them that they would skew the consensus; one target would have a disproportionate effect on the consensus. Second, although equally strong binding is observed to the clusters upstream and downstream of the transcription start point, no footprinting has been done to define the exact sites of BldD binding.

In searching for new potential BldD binding sites, certain factors were weighed more heavily. All of the confirmed BldD binding sites begin with A or G, and the following two bases are invariably GT. As well, the imperfect inverted repeat always had CAC in the center. These bases were weighted more heavily in searching for potential BldD binding sites upstream of new targets.

Table 4.1 Updated BldD binding sequences for known targets

Target name	Previously reported consensus match	Revised consensus match
bdtA (Elliot *et al.*, 2001)	AGCGA (7) TCACC	AGTGA GCACG (13) TCACC
bldD (Elliot and Leskiw, 1999)	AGTAA (7) TAACC	AGTAA (12) TCACA
bldN (1) (Elliot *et al.*, 2001)	AGTGC (16) TCTCC	AGTGC (2) GCACG
bldN (2) (Elliot *et al.*, 2001)	GGTGA (13) GCACG	
sigH (Kelemen *et al.*, 2001)	None identified	AGTGA (2) GCACG
whiG (1) (Elliot *et al.*, 2001)	AGTGA TCACC	
whiG (2) (Elliot *et al.*, 2001)	AGTCA (8) TCACG	

Previously published consensus: AGTgA (n_m) TCACc

Revised consensus: AGTGA (n_m) g/tCACg/c

Sequences that do not exactly match these criteria were identified upstream of *sigQ*, but again did not weight the search criteria as the specific sites bound by BldD have not been defined.

Identification of more BldD targets and demonstration of the exact region bound by BldD will clarify and hopefully confirm the consensus. The refined consensus presented here will help to find these targets and more accurately identify new direct targets of BldD regulation.

4.3 Weak binding of BldD to *prsI* and *arsI*

Previous studies had found very strong binding of BldD to its targets; for example, Elliot *et al.* (2001) found complete binding to the *whiG* promoter with the addition of only 1 μM BldD, and very close to complete binding of the *bldN* promoter with the same concentration of protein. It came as a surprise, then when BldD binding to *prsI* and *arsI* was considerably weaker; the K_D of BldD binding to *arsI* was 1.2 μM and binding to *prsI* was 1.0 μM. Although the BldD used in all experiments in this thesis had been purified and frozen in 1999, and protein-DNA interaction experiments did not start until late 2002, the weak binding of BldD to *prsI* and *arsI* cannot be attributed to the age of the protein. The experiments demonstrating BldD binding to these two genes were done at in the same time frame as the EMSA experiment showing BldD binding to its own promoter with a K_D of 0.4 μM, comparable to what had been seen in earlier work (Elliot, 2000). As well, the experiments showing BldD binding to the promoter of *sigQ* with a K_D of 0.3 μM were performed at the same time as the binding experiments using *prsI* and *arsI*. Finally, the *arsI* footprinting experiments were performed within weeks of the *bldD* footprinting experiments, yet the *arsI* footprints seen are not half as strong as those from the *bldD* promoter.

It is interesting that there is no correlation between the spacing between the half-sites and the amount of BldD needed to bind half of the labelled probe; examination of the BldD binding sites within *bldN* and *whiG* reveals no correlation between strength of binding and match to the consensus or length of spacing between half-sites (Table 4.2). As mentioned above, recent work (Kim *et al.*, 2006) has shown that the binding sites of *whiG* and *bldN* are bound very differently when tested separately. Table 4.2 shows that one weakly bound site has a long spacer and the other weakly bound site has no gap at all between the half sites. However, there is some correlation between the degree to which the binding site matches the consensus, and the strength of protein-DNA interaction (see Table 4.2 and Fig. 4.3). Overall, sequences with an A in the first position tend to be bound more strongly than those with a G in the first position. As well, the more weakly bound targets show almost no conservation in positions five and six of the consensus, while in the same positions the strongly bound sequences have one base predominantly or an even split between two. These differences are most clearly seen in the sequence logo representation in Figure 4.3, produced by WebLogo (Crooks *et al.*, 2004) (http://weblogo.berkeley.edu/). As a rule, the more strongly bound sequences match the overall binding consensus sequence more strongly, although the correspondence is not perfect; the binding site in the *arsI* promoter (AGTGC [12 bp] TCACG) is an excellent match to the

Table 4.2 BldD binding sequences, arranged by strength of binding

strongly bound	BldD binding sequence	weakly bound	BldD binding sequence
bldD	**AGTAA**nnnnnnnnnnnn**TCACA**	*arsI*	**AGTGC**nnnnnnnnnnnn**TCACG**
bdtA	⟶⟵ ⟵ **AGTGAGCACG**nnnnnnnnnnnnn**TCACC**	p*bldN*II	**GGTGA**nnnnnnnnnnnnn**GCACG**
p*bldN*I	**AGTGC**nn**GCACG**	*prsI* (near *prsIp2*)	**GGTCG**nnnnnnnn**ACACC**
sigH	**AGTGA**nn**GCACG**	*prsI* (near *prsIp1*)	⟶ ⟶⟶ ⟵ **GGTGG**nnn**GGTGAGGTGA**nnnnnnnnnnn**GCACC**
sigQ	many: specific binding sites not yet proven	p*whiG*I	**AGTGA TCACC**
p*whiG*II	**AGTCA**nnnnnnnn**TCACG**		

BldD binding consensus: AGTGA (n_m) g/tCACg/c

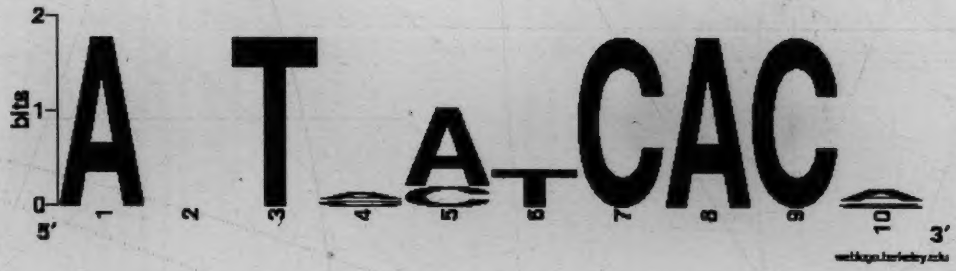
Figure 4.3

Sequence logo illustrating the sequences weakly and strongly bound by BldD. Sequence logos (generated by WebLogo at http://weblogo.berkeley.edu/) are a graphical representation of alignments of sequences, in this case the BldD binding consensus sequence. The logo is produced from statistical analysis of the level of conservation of the bases at each position within the consensus. The height of the stack at each position represents the overall level of conservation at that position, and the height of each letter within the stack represents the relative frequency of that letter at that position. Intervening sequences between the half-sites were variable in length, and were not included in the analysis. Sequence logos are designed to preserve the information lost by the traditional presentation of a consensus sequence.

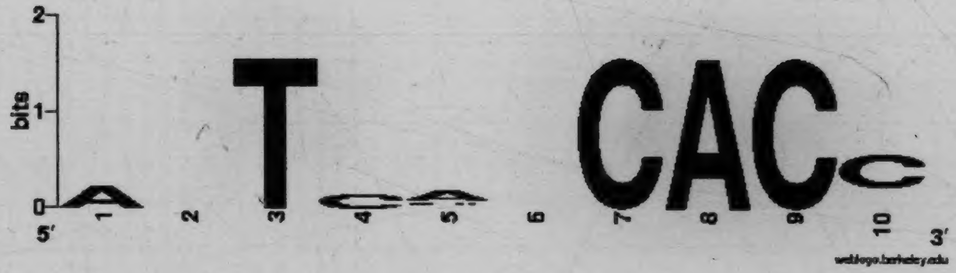
A) Consensus of strongly bound BldD targets.

B) Consensus of weakly bound BldD targets.

A


bits
2
1
0
5′
3′
1
2
3
4
5
6
7
8
9
10
A
T
A
C
T
CAC
weblogo.berkeley.edu

B


bits
2
1
0
5′
3′
1
2
3
4
5
6
7
8
9
10
A
T
C
A
CAC
C
weblogo.berkeley.edu

consensus and was bound very weakly, while the *bldD* promoter, which contains a slightly weaker match (AGTAA [12 bp] TCACA), is bound very strongly.

The consensus was derived by alignment of all of the confirmed BldD binding sites, without taking the strength of binding into consideration, however weighting the consensus does not make much difference to the final sequence. As well, it is possible that critical targets needed to define the consensus have yet to be found, and that once these are identified it will be possible to accurately predict the strength of BldD binding to a given target. Finally, given that all of the targets used to generate the consensus are repressed by BldD, more targets will have to be identified to determine if there is a separate consensus bound by BldD when activating targets.

4.4 BldD as an activator

Prior to this study, the possibility that BldD might act as a dual function regulator had been proposed (Elliot *et al.*, 2001). Both *whiG* and *bdtA* show *bldD*-dependent activation in late sporulation as well as *bldD*-dependent repression in early development in the wild type *S. coelicolor* strain 916 when compared to the congenic *bldD* point mutant 1169 (Elliot *et al.*, 2001). However, the possibility that there might be strictly activated targets of BldD was not considered. This work contains the first evidence that BldD, in addition to functioning as a repressor, may in some cases behave as an activator. This evidence came from the microarray analysis, where expression of some genes was activated instead of repressed as *bldD* expression was induced (Fig. 3.8).

Of these genes, *sigQ* was characterized further (section 3.8). Two putative induced BldD targets, *sigQ* and *sigE*, will be discussed later in this section.

BldD belongs to the XRE (xenobiotic response element) family of transcription factors. The best known members of this family are repressors. For example, lambda repressor (Maniatis and Ptashne, 1973; Maniatis *et al.*, 1975) represses the genes necessary for lytic growth of the lambda phage (Meyer *et al.*, 1980). As well, SinR (Gaur *et al.*, 1986) is a repressor in *Bacillus subtilis* that operates at a crucial checkpoint in sporulation. However, BldD would not be the first member of this family to have an activator function. ClgR in *Streptomyces lividans*, a close relative of *S. coelicolor*, has a central XRE helix-turn-helix motif. It has been shown to bind to its own promoter as well as those of *clpC1*, *clpP1* and *lon*, and to activate transcription of the *clpC1* gene and *clpP1* operon (Bellier and Mazodier, 2004). Like BldD, ClgR is involved in differentiation; overproduction of ClgR results in a delay in differentiation, although no null mutant of *clgR* has been generated (Bellier and Mazodier, 2004).

Other members of the XRE family of proteins include BcrR of *Enterococcus faecalis* (Manson *et al.*, 2004) and PigP of *Serratia* sp. ATCC 39006 (39006) (Fineran *et al.*, 2005), both of which contain N-terminal helix-turn-helix motifs. BcrR activates expression of genes involved in bacitracin resistance, notably the *bcrABD* operon, and interestingly is likely anchored in the membrane by its four transmembrane helices (Manson *et al.*, 2004). PigP is a global regulator of dual function; it activates the transcription of *pigQ, pigS* and *rap* and represses the transcription of *pigR, pigV* and *pigX*, all of which are involved in antibiotic production (Fineran *et al.*, 2005). There is precedent, then, for proteins with an XRE-type helix-turn-helix to be activators, or even to have a dual function in regulation.

sigQ was identified from the microarray analysis as a putative activated target of BldD, and the effect of *bldD* induction on *sigQ* transcript levels was more dramatic than that seen in any other targets. *sigQ* is a most unusual BldD target. Many putative BldD binding sites were found upstream of *sigQ*, clustered on either side of the transcription start point identified in the lab of Dr. Camilla Kao (N. Karoonuthaisiri, personal communication); the previously identified BldD targets have one or two BldD binding sites. Due to the position of the *sigQ* gene relative to the *afsQ* operon, the products of which are predicted to regulate *sigQ* expression (M. Hutchings, posted on ScoDB, http://streptomyces.org.uk), *sigQ* is in turn expected to play a role in the regulation of antibiotic production. Despite this, the phenotype of a *sigQ* null mutant was indistinguishable from the wild type strain (N. Karoonuthaisiri, personal communication), and the transcription pattern of *sigQ* is not consistent with this theory. In a single unrepeated experiment that addressed *sigQ* expression in solid-grown cultures (Bui, 2006), BldD regulation of *sigQ* expression showed a different pattern than had been seen for any of the previously studied BldD targets. The strains used in this experiment were the *bldD* null mutant, Δ*bldD1*, and its congenic wild type parent M600. Unlike any other targets of *bldD*, in wild type *S. coelicolor sigQ* showed high levels of expression in both early and late growth, and appeared to be repressed at the onset of aerial hyphae formation. This pattern of *sigQ* transcript abundance was abolished in the *bldD* null mutant where *sigQ* transcripts were constitutively expressed. This transcription pattern argues against *sigQ* being involved directly in antibiotic production because *sigQ* expression was high at the earliest stage of growth (prior to antibiotic production) and late in sporulation (well after antibiotic production has been established), and was completely absent at the phase where antibiotic production begins. SigQ protein levels were not addressed in this work, so no comment can be made about possible post-translational regulation, or stability of the protein in the absence of transcript. However, from both mutational analysis and transcriptional regulation, the role of *sigQ* in development does not appear to be linked to antibiotic production.

From evidence presented in Karoonuthaisiri *et al.* (2005), *sigQ* seems to have a role in vegetative growth in liquid culture. Overall, during vegetative growth genes involved in primary metabolism are expressed at a significantly higher level than regulatory genes. Certain regulatory genes, including *hrdB* (the principal vegetative sigma factor of *S. coelicolor*) and *sigQ*, form an exception to this rule and are expressed four to eight times more strongly than the average expression levels (Karoonuthaisiri *et al.*, 2005). It seems odd that *bldD*, a gene so important for development, would regulate a gene involved in vegetative growth. One possible reason for *bldD*-dependent repression of *sigQ* during aerial hyphae formation could be to prevent SigQ-directed transcription of genes that would interfere with the change in growth phase. An alternative theory could be that BldD is active only during very early growth, when the strongest repression of the known targets is seen. This would mean that the absence of *sigQ* expression during aerial hyphae formation is because of a loss of *bldD*-dependent activation, rather than active repression. However, this doesn't explain the increased abundance of *sigQ* transcripts late in growth, nor the reduced constitutive and *bldD*-independent level of *sigQ* expression seen in the *bldD* mutant strain. This model is therefore less likely to be true. Confirmation of the expression pattern of *sigQ* from growth on solid media is important, especially given the unusual pattern. It would also be very interesting to see if the higher than average expression of *sigQ* during vegetative growth in liquid culture described by Karoonuthaisiri *et al.* (2005) is dependent on *bldD*.

As seen in section 3.8 of Results, there are many putative BldD binding sites upstream of *sigQ*, and these sites are clustered into two groups. Although binding to both clusters, individually and together, has been demonstrated, no footprinting experiments have been performed to identify which particular site(s) are bound. This never before seen pattern of putative BldD binding sites is most likely part of the unusual regulation of *sigQ* by BldD. An intriguing possibility is that BldD binds to one set of binding sites to repress expression and the other to induce expression. Interestingly, although *bldD* might be necessary for both repression and induction of *sigQ* in solid culture, induction of *bldD* in liquid culture of the *bldD* mutant strain (and presumably therefore production of large quantities of BldD protein) resulted solely in activation of *sigQ* (Fig. 3.9, E and F) (Bui, 2006). This may mean that in footprinting experiments, where artificially large concentrations of BldD protein are used, only those sites bound to activate *sigQ* will be visualized. However, this simplistic model assumes that the regulation of *sigQ* by BldD is concentration-dependent and not a result of cofactors or conditions found in the culture. Because the BldD protein used for *in vitro* experiments was purified from *E. coli*, not *S. coelicolor*, any cofactors or binding partners that might be present in *S. coelicolor* are absent and so no conclusions can be drawn regarding cofactor involvement. As well, the chromatin immunoprecipitation (ChIP) experiments done in this study could not have resolved which sites are bound at different points in development because the smallest fragments generated were still close to 500 bp. The resolution needed for determination of which sites are bound would require fragments of less than 100 bp. It is possible that once the mechanism of switching from repressor to activator function of BldD is known, then the specific location of BldD binding to the *sigQ* will become clearer.

In retrospect, one of the most interesting, although not studied in the context of this thesis, potential BldD targets found by microarray analysis was *sigE*. *sigE* encodes an ECF sigma factor required for normal cell envelope stability in *S. coelicolor* (Paget *et al.*, 1999a). It seems very odd that BldD would upregulate *sigE* transcription. There is a significant level of basal expression from the *sigE* promoter, proposed to be due to normal changes in the cell envelope (Hong *et al.*, 2002), besides a demonstrated dependence on *cseB* for any level of expression (Paget *et al.*, 1999b). *sigE* expression is activated by the CseBC two component regulatory system, and has been recently shown to be enhanced in the absence of *cseA* which encodes a lipoprotein proposed to interact with the sensor kinase CseC or to stabilize the cell membrane (Hutchings *et al.*, 2006). Expression of *sigE* is induced by a wide variety of structurally unrelated antibiotics, all of which interfere with cell wall biosynthesis, as well as lysozyme, which digests peptidoglycan (Hong *et al.*, 2002). *sigE* null mutants require high levels of divalent cations (Ca^{2+} or Mg^{2+}), in the absence of which they form crenellated colonies that sporulate poorly and overproduce actinorhodin (Paget *et al.*, 1999a). None of these factors would indicate any role for BldD regulation, and the complete dependence of *sigE* expression on *cseB* would argue for a negative regulatory role for BldD, not a positive one. Despite this, the evidence from the microarray experiment showed induction of not only *sigE* but also the co-transcribed genes *cseB* and *cseC* (Fig. C.1, C and D). It is possible that *bldD*-dependent activation of *sigE* transcription is specific to a particular stimulus rather than developmental phase. This would make sense, given that there is no significant difference between the putative BldD binding site upstream of *sigE* and those in front of targets shown to be repressed. The stimulus could result in a short-term small molecule effector for BldD, with the binding and activation only occurring under the specific conditions where BldD is needed as an activator. Alternately, the stimulus could be specific to a particular stage in development. Erection of aerial hyphae involves not only interaction with a completely new environment but also production of a whole range of new proteins that coat the cell surface (Claessen *et al.*, 2003; Elliot *et al.*, 2003a). It is perhaps possible that BldD induction of the *sigE* operon could play a pre-emptive role in dealing with the stresses, although much work would need to be done to prove the link, including verification that *sigE* is a genuine BldD target.

The position of the putative BldD binding site upstream of *sigE* is somewhat unusual. An activator would be expected to bind upstream of the transcription start point, where it would be able to recruit the RNA polymerase to the promoter without at the same time physically blocking its progress. Thus, the placement of the site between the transcription start point and the translational start codon (48 bp downstream of the transcription start point) is more consistent with the location of repressor binding. This placement, however unusual, is not without precedent. In *Streptomyces griseus*, the A-factor dependent transcriptional activator AdpA binds to the promoter of the *bldN* homologue *adsA* between +7 and +46 relative the transcription start point (Yamazaki *et al.*, 2000), a position directly overlapping the binding site for RNA polymerase. Therefore, the placement of the BldD site, by itself, is unusual but does not eliminate the possibility that *sigE* is an activated target of BldD.

Of the two genes discussed in this section, the second (*sigE*) appears to be solely activated by BldD but the first (*sigQ*) appears to be both activated and repressed by BldD. Possible mechanisms for this switch are discussed below.

4.5 BldD as a dual-function regulator

The nature of the mechanism that converts BldD from repressor to activator is an important question. How is it that under some circumstances BldD represses its targets, yet under others it activates them? As well, is this switching mechanism the same as the means by which BldD repression is relieved? Phosphorylation on Tyr62 and binding to a protein cofactor have both been proposed as means of regulation of BldD activity (Elliot *et al.*, 1998; Elliot *et* *al.*, 2001). The former has been removed from consideration because *in vitro* phosphorylation of BldD was not possible, despite multiple attempts (Elliot, 2000) and the most likely site of phosphorylation (Tyr62) has been shown to be within the hydrophobic core of the protein (Kim *et al.*, 2006) and thus inaccessible to hydrophilic modifications. The latter option is an interesting one. There is precedent within the XRE family for binding of a partner protein for modulation of regulator activity. SinR of *Bacillus subtilis* is a developmental regulator that represses key genes required for sporulation, and which binds its targets as a homotetramer (Lewis *et al.*, 1998). Repression is relieved by binding of SinI to form heterodimers with no DNA-binding activity (Lewis *et al.*, 1998). This binding is obviously not the sort to be expected for a change of function, but rather a removal of function. Unfortunately, an intensive search has failed to find any evidence of a protein binding partner for BldD (Bui, 2006), making this model seem unlikely.

Positive or negative regulatory activity of DNA-binding regulatory proteins can also be determined by concentration. AdpA of *Streptomyces griseus* (of the AraC/XylS family) is a key regulator necessary for activation of genes required for development and secondary metabolism (Ohnishi *et al.*, 2005). AdpA is proposed to self-regulate by binding sequentially to three sites in its own promoter in a concentration-dependent manner. When the cellular concentration is sufficient, AdpA dimers interact to bend the DNA, occluding the promoter and preventing transcription (Kato *et al.*, 2005). Although this type of model could apply to a wild type strain such as M600, where the BldD levels decrease dramatically in late growth (Bui, 2006), the decrease in protein levels happens well after BldD repression of the known targets is relieved. As well, the wild type strain J1501 showed no such reduction in protein levels (Elliot, 2000). In addition, *sigQ* induction is seen in early growth where all other BldD targets are repressed. A model to explain BldD regulation would have to explain its function in all wild type strains, and should describe all targets.

The question of how BldD repression is relieved, or how BldD is converted from a repressor to an activator, would be rendered much simpler if there were enough targets of each type (activated, repressed or both) for a pattern to be seen in the sequences bound by BldD. The targets of BldD described in Figure 4.1 are repressed in early growth (*arsI, prsI*), repressed throughout growth (*bldD, bldN*), or repressed then potentially activated (*bdtA, whiG*); the expression of *sigH* has not been examined in different stages of development. As well, from a single solid-culture experiment *sigQ* appeared to be alternately activated and repressed. Although numerous potential BldD targets have been identified by this work, only the transcriptional regulation of *prsI*, *arsI* and *sigQ* has been examined, and of these *sigQ* has not been thoroughly examined (this work and Bui, 2006). More work is needed to determine if the pattern of *sigQ* transcription seen is reproducible. In short, one likely activated target, with the exact location of BldD binding unknown, is not enough to draw conclusions from.

Although there are not enough characterized targets to determine if the exact nature of the BldD binding sequence determines whether BldD binds as an activator or a repressor, some conclusions can be drawn regarding the position of the binding relative to the transcription start point of target genes. If the placement of BldD binding relative to the transcription start point was the factor that determined the activity (positive or negative regulator), then there should be a pattern in the placement of the BldD binding sites of targets known to be repressed. As seen in Figure 4.1, the distance between BldD binding sites and the transcription start point of strictly repressed targets ranges from approximately 20 bp upstream to 8 bp downstream, all consistent with repression. However, among the strictly repressed targets are found *bldD* and *arsI*, each of which have one BldD binding site, while *bldN* which has two. *whiG*, the target with one BldD binding site at a position most suitable for an activator, -120 to -110 relative to the tsp, is clearly repressed in early development and its apparent *bldD*-dependent activation is more easily attributed to factors unrelated to direct BldD regulation. From these data, it therefore seems unlikely that the same active form of BldD acts both as activator and repressor, depending simply on the position of the binding sites to determine its function. More likely, a combination of factors acts together to determine the function of BldD as an activator or a repressor.

The most likely mechanism for modifying the activity of BldD is binding by a small cofactor. AraC of *E. coli*, a well known example of this type of regulation, binds the promoter of the *araBAD* operon constantly. It serves as a repressor in the absence of arabinose and as an activator when bound to arabinose (Carra and Schleif, 1993). In the case of BldD regulation, a cofactor might be present in late development (when BldD activates *bdtA* and *whiG* transcription) but not during early development, prior to aerial hyphae formation (when BldD represses *bldN*, *whiG* and itself, among others). Overexpression of *bldN* and *sigH* in late development in a *bldD* mutant (Elliot *et al.*, 2001; Kelemen *et al.*, 2001) could be addressed within this model. BldD restricts the expression of *sigH* to the aerial hyphae, meaning that it represses *sigH* expression in the substrate mycelium. In the absence of BldD, *sigH* is expressed ectopically in the substrate mycelium (Kelemen *et al.*, 2001). This ectopic expression is likely the cause of the overexpression seen in the *bldD* mutant. Spatial regulation of *bldN* has not been tested, but if it, too, is restricted to a particular tissue by BldD then the same explanation would hold. Early overexpression of *sigQ* is not explained by this model, but the unusual gene structure in the *sigQ* promoter may allow a mode of regulation unique to this gene. Experiments to address a small cofactor binding to BldD have not been performed and would be difficult without at least an idea what the cofactor might be. There is a plethora of putative secreted proteins encoded in the *S. coelicolor* genome (Bentley *et al.*, 2002). Perhaps one of these is imported when it reaches a threshold extracellular concentration to serve as a signalling molecule within the cell, converting BldD to its new function. If this concentration was only reached in certain tissues, it could even affect spatial regulation by BldD.

4.6 Conclusion

Overall, this work has succeeded in its goal of finding new targets of BldD regulation. These targets have been both expected (repressed by BldD) and surprising (activated by BldD). It is still not known how *bldD* is involved in regulation of secondary metabolite production; mutants of *sigQ*, which was proposed to be a link between *bldD* and antibiotic production, showed no defect in antibiotic production. Still, the unique nature of the sequence upstream of *sigQ* and the potentially cyclic nature of the BldD regulation of its transcription makes it an interesting model with which to study how BldD is converted from repressor to activator.

As a result of work done in this thesis, the consensus sequence for BldD binding has undergone revision. Although the expected spacer has not been changed, the sequence has been shown to be somewhat less defined than previously reported. This emphasizes the necessity for methods of finding new BldD targets that do not rely exclusively on sequence.

There are many more genes identified by the microarray analysis that were not confirmed or tested in any way. Those chosen for further analysis in this work have raised many questions of their own, but in the process added to the body of knowledge about *Streptomyces* development. Further analysis of other putative targets should help form a clearer picture of the role of BldD.

Bibliography

Ainsa, J.A., Parry, H.D., and Chater, K.F. (1999) A response regulator-like protein that functions at an intermediate stage of sporulation in *Streptomyces coelicolor* A3(2). *Mol Microbiol* **34**: 607-619.

Ainsa, J.A., Ryding, N.J., Hartley, N., Findlay, K.C., Bruton, C.J., and Chater, K.F. (2000) WhiA, a protein of unknown function conserved among Gram-positive bacteria, is essential for sporulation in *Streptomyces coelicolor* A3(2). *J Bacteriol* **182**: 5470-5478.

Alper, S., Duncan, L., and Losick, R. (1994) An adenosine nucleotide switch controlling the activity of a cell type-specific transcription factor in *B. subtilis*. *Cell* **77**: 195-205.

Anderson, T.B., Brian, P., and Champness, W.C. (2001) Genetic and transcriptional analysis of *absA*, an antibiotic gene cluster-linked two-component system that regulates multiple antibiotics in *Streptomyces coelicolor*. *Mol Microbiol* **39**: 553-566.

Arias, P., Fernandez-Moreno, M.A., and Malpartida, F. (1999) Characterization of the pathway-specific positive transcriptional regulator for actinorhodin biosynthesis in Streptomyces coelicolor A3(2) as a DNA-binding protein. *J Bacteriol* **181**: 6958-6968.

Bellier, A., and Mazodier, P. (2004) ClgR, a novel regulator of clp and lon expression in Streptomyces. *J Bacteriol* **186**: 3238-3248.

Bentley, S.D., Chater, K.F., Cerdeno-Tarraga, A.M., Challis, G.L., Thomson, N.R., James, K.D., Harris, D.E., Quail, M.A., Kieser, H., Harper, D., Bateman, A., Brown, S., Chandra, G., Chen, C.W., Collins, M., Cronin, A., Fraser, A., Goble, A., Hidalgo, J., Hornsby, T., Howarth, S., Huang, C.H., Kieser, T., Larke, L., Murphy, L., Oliver, K., O'Neil, S., Rabbinowitsch, E., Rajandream, M.A., Rutherford, K., Rutter, S., Seeger, K., Saunders, D., Sharp, S., Squares, R., Squares, S., Taylor, K., Warren, T., Wietzorrek, A., Woodward, J., Barrell, B.G., Parkhill, J., and Hopwood, D.A. (2002) Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* **417**: 141-147.

Bertani, G. (2004) Lysogeny at mid-twentieth century: P1, P2, and other experimental systems. *J Bacteriol* **186**: 595-600.

Bibb, M. (1996) 1995 Colworth Prize Lecture. The regulation of antibiotic production in *Streptomyces coelicolor* A3(2). *Microbiology* **142**: 1335-1344.

Bibb, M.J., Molle, V., and Buttner, M.J. (2000) σ^{BldN}, an extracytoplasmic function RNA polymerase sigma factor required for aerial mycelium formation in *Streptomyces coelicolor* A3(2). *J Bacteriol* **182**: 4606-4616.

Bibb, M.J., and Buttner, M.J. (2003) The *Streptomyces coelicolor* Developmental Transcription Factor σ^{BldN} Is Synthesized as a Proprotein. *J Bacteriol* **185**: 2338-2345.

Bierman, M., Logan, R., O'Brien, K., Seno, E.T., Rao, R.N., and Schoner, B.E. (1992) Plasmid cloning vectors for the conjugal transfer of DNA from Escherichia coli to Streptomyces spp. *Gene* **116**: 43-49.

Bignell, D.R., Warawa, J.L., Strap, J.L., Chater, K.F., and Leskiw, B.K. (2000) Study of the *bldG* locus suggests that an anti-anti-sigma factor and an anti-sigma factor may be involved in *Streptomyces coelicolor* antibiotic production and sporulation. *Microbiology* **146**: 2161-2173.

Bignell, D.R., Lau, L.H., Colvin, K.R., and Leskiw, B.K. (2003) The putative anti-anti-sigma factor BldG is post-translationally modified by phosphorylation in *Streptomyces coelicolor. Submitted.*

Bignell, D.R., Tahlan, K., Colvin, K.R., Jensen, S.E., and Leskiw, B.K. (2005) Expression of ccaR, encoding the positive activator of cephamycin C and clavulanic acid production in Streptomyces clavuligerus, is dependent on bldG. *Antimicrob Agents Chemother* **49**: 1529-1541.

Bignell, D.R.D. (2003) Characterization of the bldG locus in Streptomyces coelicolor. In *Biological Sciences* Edmonton: University of Alberta, pp. 338.

Bishop, A., Fielding, S., Dyson, P., and Herron, P. (2004) Systematic insertional mutagenesis of a streptomycete genome: a link between osmoadaptation and antibiotic production. *Genome Res* **14**: 893-900.

Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**: 248-254.

Brian, P., Riggle, P.J., Santos, R.A., and Champness, W.C. (1996) Global negative regulation of *Streptomyces coelicolor* antibiotic synthesis mediated by an *absA*-encoded putative signal transduction system. *J Bacteriol* **178**: 3221-3231.

Bui, L. (2006) Investigating post-translational control of BldD in *Streptomyces coelicolor*. In *Biological Sciences* Edmonton: University of Alberta.

Carra, J.H., and Schleif, R.F. (1993) Variation of half-site organization and DNA looping by AraC protein. *Embo J* **12**: 35-44.

Chaconas, G., and van de Sande, J.H. (1980) 5'-32P labeling of RNA and DNA restriction fragments. *Methods Enzymol* **65**: 75-85.

Challis, G.L., and Hopwood, D.A. (2003) Synergy and contingency as driving forces for the evolution of multiple secondary metabolite production by Streptomyces species. *Proc Natl Acad Sci U S A* **100 Suppl 2**: 14555-14561.

Champness, W.C. (1988) New loci required for *Streptomyces coelicolor* morphological and physiological differentiation. *J Bacteriol* **170**: 1168-1174.

Chater, K.F. (1972) A morphological and genetic mapping study of white colony mutants of *Streptomyces coelicolor*. *J Gen Microbiol* **72**: 9-28.

Chater, K.F. (1975) Construction and phenotypes of double sporulation deficient mutants in Streptomyces coelicolor A3(2). *J Gen Microbiol* **87**: 312-325.

Chater, K.F., Bruton, C.J., King, A.A., and Suarez, J.E. (1982) The expression of Streptomyces and Escherichia coli drug-resistance determinants cloned into the Streptomyces phage phi C31. *Gene* **19**: 21-32.

Chater, K.F. (1998) Taking a genetic scalpel to the *Streptomyces* colony. *Microbiology* **144**: 1465-1478.

Chater, K.F. (2001) Regulation of sporulation in *Streptomyces coelicolor* A3(2): a checkpoint multiplex? *Curr Opin Microbiol* **4**: 667-673.

Chater, K.F. (2006) Streptomyces inside-out: a new perspective on the bacteria that provide us with antibiotics. *Philos Trans R Soc Lond B Biol Sci* **361**: 761-768.

Chater, K.F., and Chandra, G. (2006) The evolution of development in Streptomyces analysed by genome comparisons. *FEMS Microbiol Rev* **30**: 651-672.

Chong, P.P., Podmore, S.M., Kieser, H.M., Redenbach, M., Turgay, K., Marahiel, M., Hopwood, D.A., and Smith, C.P. (1998) Physical identification of a chromosomal locus encoding biosynthetic genes for the lipopeptide calcium-dependent antibiotic (CDA) of *Streptomyces coelicolor* A3(2). *Microbiology* **144**: 193-199.

Claessen, D., Rink, R., de Jong, W., Siebring, J., de Vreugd, P., Boersma, F.G., Dijkhuizen, L., and Wosten, H.A. (2003) A novel class of secreted hydrophobic proteins is involved in aerial hyphae formation in Streptomyces coelicolor by forming amyloid-like fibrils. *Genes Dev* **17**: 1714-1726.

Claessen, D., Stokroos, I., Deelstra, H.J., Penninga, N.A., Bormann, C., Salas, J.A., Dijkhuizen, L., and Wosten, H.A. (2004) The formation of the rodlet layer of streptomycetes is the result of the interplay between rodlins and chaplins. *Mol Microbiol* **53**: 433-443.

Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E. (2004) WebLogo: a sequence logo generator. *Genome Res* **14**: 1188-1190.

Davis, N.K., and Chater, K.F. (1992) The *Streptomyces coelicolor whiB* gene encodes a small transcription factor-like protein dispensable for growth but essential for sporulation. *Mol Gen Genet* **232**: 351-358.

Eccleston, M., Ali, R.A., Seyler, R., Westpheling, J., and Nodwell, J. (2002) Structural and genetic analysis of the BldB protein of *Streptomyces coelicolor*. *J Bacteriol* **184**: 4270-4276.

Eccleston, M., Willems, A., Beveridge, A., and Nodwell, J.R. (2006) Critical residues and novel effects of overexpression of the Streptomyces coelicolor developmental protein BldB: evidence for a critical interacting partner. *J Bacteriol* **188**: 8189-8195.

Elliot, M., Damji, F., Passantino, R., Chater, K., and Leskiw, B. (1998) The *bldD* gene of *Streptomyces coelicolor* A3(2): a regulatory gene involved in morphogenesis and antibiotic production. *J Bacteriol* **180**: 1549-1555.

Elliot, M.A., and Leskiw, B.K. (1999) The BldD protein from *Streptomyces coelicolor* is a DNA-binding protein. *J Bacteriol* **181**: 6832-6835.

Elliot, M.A. (2000) Characterization of BldD and its targets in Streptomyces coelicolor. Edmonton, AB: University of Alberta.

Elliot, M.A., Bibb, M.J., Buttner, M.J., and Leskiw, B.K. (2001) BldD is a direct regulator of key developmental genes in *Streptomyces coelicolor* A3(2). *Mol Microbiol* **40**: 257-269.

Elliot, M.A., Karoonuthaisiri, N., Huang, J., Bibb, M.J., Cohen, S.N., Kao, C.M., and Buttner, M.J. (2003a) The chaplins: a family of hydrophobic cell-surface proteins involved in aerial mycelium formation in *Streptomyces coelicolor*. *Submitted*.

Elliot, M.A., Locke, T.R., Galibois, C.M., and Leskiw, B.K. (2003b) BldD from *Streptomyces coelicolor* is a non-essential global regulator that binds DNA as a dimer. *Submitted.*

Feinberg, A.P., and Vogelstein, B. (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal Biochem* **132**: 6-13.

Fernandez-Moreno, M.A., Caballero, J.L., Hopwood, D.A., and Malpartida, F. (1991) The *act* cluster contains regulatory and antibiotic export genes, direct targets for translational control by the *bldA* tRNA gene of *Streptomyces*. *Cell* **66**: 769-780.

Fineran, P.C., Slater, H., Everson, L., Hughes, K., and Salmond, G.P. (2005) Biosynthesis of tripyrrole and beta-lactam secondary metabolites in Serratia: integration of quorum sensing with multiple new regulatory components in the control of prodigiosin and carbapenem antibiotic production. *Mol Microbiol* **56**: 1495-1517.

Fisher, S.H., and Sonenshein, A.L. (1991) Control of carbon and nitrogen metabolism in Bacillus subtilis. *Annu Rev Microbiol* **45**: 107-135.

Flardh, K., Findlay, K.C., and Chater, K.F. (1999) Association of early sporulation genes with suggested developmental decision points in *Streptomyces coelicolor* A3(2). *Microbiology* **145**: 2229-2243.

Flardh, K., Leibovitz, E., Buttner, M.J., and Chater, K.F. (2000) Generation of a non-sporulating strain of Streptomyces coelicolor A3(2) by the manipulation of a developmentally controlled ftsZ promoter. *Mol Microbiol* **38**: 737-749.

Gaur, N.K., Dubnau, E., and Smith, I. (1986) Characterization of a cloned Bacillus subtilis gene that inhibits sporulation in multiple copies. *J Bacteriol* **168**: 860-869.

Gehring, A.M., Nodwell, J.R., Beverley, S.M., and Losick, R. (2000) Genomewide insertional mutagenesis in *Streptomyces coelicolor* reveals additional genes involved in morphological differentiation. *Proc Natl Acad Sci U S A* **97**: 9642-9647.

Gehring, A.M., Yoo, N.J., and Losick, R. (2001) RNA polymerase sigma factor that blocks morphological differentiation by *Streptomyces coelicolor*. *J Bacteriol* **183**: 5991-5996.

Gust, B., Challis, G.L., Fowler, K., Kieser, T., and Chater, K.F. (2003) PCR-targeted Streptomyces gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. *Proc Natl Acad Sci U S A* **100**: 1541-1546.

Guthrie, E.P., Flaxman, C.S., White, J., Hodgson, D.A., Bibb, M.J., and Chater, K.F. (1998) A response-regulator-like activator of antibiotic synthesis from *Streptomyces coelicolor* A3(2) with an amino-terminal domain that lacks a phosphorylation pocket. *Microbiology* **144**: 727-738.

Hakenbeck, R., and Stock, J.B. (1996) Analysis of two-component signal transduction systems involved in transcriptional regulation. *Methods Enzymol* **273**: 281-300.

Hanahan, D. (1983) Studies on transformation of Escherichia coli with plasmids. *J Mol Biol* **166**: 557-580.

Hobbs, G., Obanye, A.I., Petty, J., Mason, J.C., Barratt, E., Gardner, D.C., Flett, F., Smith, C.P., Broda, P., and Oliver, S.G. (1992) An integrated approach to studying regulation of production of the antibiotic methylenomycin by Streptomyces coelicolor A3(2). *J Bacteriol* **174**: 1487-1494.

Hong, H.J., Paget, M.S., and Buttner, M.J. (2002) A signal transduction system in *Streptomyces coelicolor* that activates the expression of a putative cell wall glycan operon in response to vancomycin and other cell wall-specific antibiotics. *Mol Microbiol* **44**: 1199-1211.

Hopwood, D.A., Bibb, M.J., Chater, K.F., Kieser, T., Bruton, C.J., Kieser, H.M., Lydiate, D.J., Smith, C.P., Ward, J.M., and Schrempf, H. (1985) *Genetic Manipulation of Streptomyces, a laboratory manual.* Norwich: The John Innes Foundation.

Hopwood, D.A., Chater, K.F., and Bibb, M.J. (1994) Genetics and biochemistry of antibiotic production in *Streptomyces coelicolor* A3(2). In *Genetics and biochemistry of antibiotic production*. Vinning, L.C. and Stuttard, C. (eds): Butterwerth-Heinemann, pp. 65-102.

Horinouchi, S. (2003) AfsR as an integrator of signals that are sensed by multiple serine/threonine kinases in Streptomyces coelicolor A3(2). *J Ind Microbiol Biotechnol* **30**: 462-467.

Huang, J., Shi, J., Molle, V., Sohlberg, B., Weaver, D., Bibb, M.J., Karoonuthaisiri, N., Lih, C.J., Kao, C.M., Buttner, M.J., and Cohen, S.N. (2005) Cross-regulation among disparate antibiotic biosynthetic pathways of Streptomyces coelicolor. *Mol Microbiol* **58**: 1276-1287.

Hudson, M.E., and Nodwell, J.R. (2004) Dimerization of the RamC morphogenetic protein of Streptomyces coelicolor. *J Bacteriol* **186**: 1330-1336.

Hughes, K.T., and Mathee, K. (1998) The anti-sigma factors. *Annu Rev Microbiol* **52**: 231-286.

Hunt, A.C., Servin-Gonzalez, L., Kelemen, G.H., and Buttner, M.J. (2005) The bldC developmental locus of Streptomyces coelicolor encodes a member of a family of small DNA-binding proteins related to the DNA-binding domains of the MerR family. *J Bacteriol* **187**: 716-728.

Hutchings, M.I., Hong, H.J., Leibovitz, E., Sutcliffe, I.C., and Buttner, M.J. (2006) The sigma(E) cell envelope stress response of Streptomyces coelicolor is influenced by a novel lipoprotein, CseA. *J Bacteriol* **188**: 7222-7229.

Ikeda, H., Ishikawa, J., Hanamoto, A., Shinose, M., Kikuchi, H., Shiba, T., Sakaki, Y., Hattori, M., and Omura, S. (2003) Complete genome sequence and comparative analysis of the industrial microorganism Streptomyces avermitilis. *Nat Biotechnol* **21**: 526-531.

Ishizuka, H., Horinouchi, S., Kieser, H.M., Hopwood, D.A., and Beppu, T. (1992) A putative two-component regulatory system involved in secondary metabolism in *Streptomyces* spp. *J Bacteriol* **174**: 7585-7594.

Jakimowicz, D., Chater, K., and Zakrzewska-Czerwinska, J. (2002) The ParB protein of Streptomyces coelicolor A3(2) recognizes a cluster of parS sequences within the origin-proximal region of the linear chromosome. *Mol Microbiol* **45**: 1365-1377.

Jakimowicz, P., Cheesman, M.R., Bishai, W.R., Chater, K.F., Thomson, A.J., and Buttner, M.J. (2005) Evidence that the Streptomyces developmental protein WhiD, a member of the WhiB family, binds a [4Fe-4S] cluster. *J Biol Chem* **280**: 8309-8315.

Janssen, G.R., and Bibb, M.J. (1993) Derivatives of pUC18 that have BglII sites flanking a modified multiple cloning site and that retain the ability to identify recombinant clones by visual screening of Escherichia coli colonies. *Gene* **124**: 133-134.

Karlinsey, J.E., Tanaka, S., Bettenworth, V., Yamaguchi, S., Boos, W., Aizawa, S.I., and Hughes, K.T. (2000) Completion of the hook-basal body complex of the Salmonella typhimurium flagellum is coupled to FlgM secretion and fliC transcription. *Mol Microbiol* **37**: 1220-1231.

Karoonuthaisiri, N., Weaver, D., Huang, J., Cohen, S.N., and Kao, C.M. (2005) Regional organization of gene expression in Streptomyces coelicolor. *Gene* **353**: 53-66.

Kato, J.Y., Ohnishi, Y., and Horinouchi, S. (2005) Autorepression of AdpA of the AraC/XylS family, a key transcriptional activator in the A-factor regulatory cascade in Streptomyces griseus. *J Mol Biol* **350**: 12-26.

Keijser, B.J., van Wezel, G.P., Canters, G.W., Kieser, T., and Vijgenboom, E. (2000) The *ram*-dependence of *Streptomyces lividans* differentiation is bypassed by copper. *J Mol Microbiol Biotechnol* **2**: 565-574.

Keijser, B.J., Noens, E.E., Kraal, B., Koerten, H.K., and van Wezel, G.P. (2003) The Streptomyces coelicolor ssgB gene is required for early stages of sporulation. *FEMS Microbiol Lett* **225**: 59-67.

Kelemen, G.H., Brown, G.L., Kormanec, J., Potuckova, L., Chater, K.F., and Buttner, M.J. (1996) The positions of the sigma-factor genes, *whiG* and *sigF*, in the hierarchy controlling the development of spore chains in the aerial hyphae of *Streptomyces coelicolor* A3(2). *Mol Microbiol* **21**: 593-603.

Kelemen, G.H., and Buttner, M.J. (1998) Initiation of aerial mycelium formation in *Streptomyces*. *Curr Opin Microbiol* **1**: 656-662.

Kelemen, G.H., Viollier, P.H., Tenor, J., Marri, L., Buttner, M.J., and Thompson, C.J. (2001) A connection between stress and development in the multicellular prokaryote *Streptomyces coelicolor* A3(2). *Mol Microbiol* **40**: 804-814.

Kieser, T., Bibb, M.J., Buttner, M.J., Chater, K.F., and Hopwood, D.A. (2000) *Practical Streptomyces Genetics*. Norwich: The John Innes Foundation.

Kim, D.W., Chater, K., Lee, K.J., and Hesketh, A. (2005) Changes in the extracellular proteome caused by the absence of the bldA gene product, a developmentally significant tRNA, reveal a new target for the pleiotropic regulator AdpA in Streptomyces coelicolor. *J Bacteriol* **187**: 2957-2966.

Kim, I.K., Lee, C.J., Kim, M.K., Kim, J.M., Kim, J.H., Yim, H.S., Cha, S.S., and Kang, S.O. (2006) Crystal structure of the DNA-binding domain of BldD, a central regulator of aerial mycelium formation in Streptomyces coelicolor A3(2). *Mol Microbiol* **60**: 1179-1193.

Kodani, S., Hudson, M.E., Durrant, M.C., Buttner, M.J., Nodwell, J.R., and Willey, J.M. (2004) The SapB morphogen is a lantibiotic-like peptide derived from the product of the developmental gene ramS in Streptomyces coelicolor. *Proc Natl Acad Sci U S A* **101**: 11448-11453.

Korn-Wendisch, F., and Kutzner, H.J. (1992) The Family Streptomycetaceae. In *The Prokaryotes: a handbook on the biology of bacteria: ecophysiology, isolation, identification, applications*. Balows, A. (ed). New York: Springer-Verlag, pp. 921-980.

Kuo, S., Zhang, S., Woodbury, R.L., and Haldenwang, W.G. (2004) Associations between Bacillus subtilis sigmaB regulators in cell extracts. *Microbiology* **150**: 4125-4136.

Larson, J.L., and Hershberger, C.L. (1986) The minimal replicon of a streptomycete plasmid produces an ultrahigh level of plasmid DNA. *Plasmid* **15**: 199-209.

Lee, E.J., Karoonuthaisiri, N., Kim, H.S., Park, J.H., Cha, C.J., Kao, C.M., and Roe, J.H. (2005) A master regulator sigmaB governs osmotic and oxidative response as well as differentiation via a network of sigma factors in Streptomyces coelicolor. *Mol Microbiol* **57**: 1252-1264.

Leskiw, B.K., Mah, R., Lawlor, E.J., and Chater, K.F. (1993) Accumulation of *bldA*-specified tRNA is temporally regulated in *Streptomyces coelicolor* A3(2). *J Bacteriol* **175**: 1995-2005.

Leskiw, B.K., and Mah, R. (1995) The *bldA*-encoded tRNA is poorly expressed in the *bldI* mutant of *Streptomyces coelicolor* A3(2). *Microbiology* **141**: 1921-1926.

Lewis, R.J., Brannigan, J.A., Offen, W.A., Smith, I., and Wilkinson, A.J. (1998) An evolutionary link between sporulation and prophage induction in the structure of a repressor:anti-repressor complex. *J Mol Biol* **283**: 907-912.

Lomovskaya, N.D., Chater, K.F., and Mkrtumian, N.M. (1980) Genetics and molecular biology of Streptomyces bacteriophages. *Microbiol Rev* **44**: 206-229.

Ma, H., and Kendall, K. (1994) Cloning and analysis of a gene cluster from *Streptomyces coelicolor* that causes accelerated aerial mycelium formation in *Streptomyces lividans*. *J Bacteriol* **176**: 3800-3811.

MacNeil, D.J., Gewain, K.M., Ruby, C.L., Dezeny, G., Gibbons, P.H., and MacNeil, T. (1992) Analysis of Streptomyces avermitilis genes required for avermectin biosynthesis utilizing a novel integration vector. *Gene* **111**: 61-68.

Maniatis, T., and Ptashne, M. (1973) Multiple repressor binding at the operators in bacteriophage lambda. *Proc Natl Acad Sci U S A* **70**: 1531-1535.

Maniatis, T., Ptashne, M., Backman, K., Kield, D., Flashman, S., Jeffrey, A., and Maurer, R. (1975) Recognition sequences of repressor and polymerase in the operators of bacteriophage lambda. *Cell* **5**: 109-113.

Manson, J.M., Keis, S., Smith, J.M., and Cook, G.M. (2004) Acquired bacitracin resistance in Enterococcus faecalis is mediated by an ABC transporter and a novel regulatory protein, BcrR. *Antimicrob Agents Chemother* **48**: 3743-3748.

Manteca, A., Mader, U., Connolly, B.A., and Sanchez, J. (2006) A proteomic analysis of Streptomyces coelicolor programmed cell death. *Proteomics* **6**: 6008-6022.

McCormick, J.R., Su, E.P., Driks, A., and Losick, R. (1994) Growth and viability of Streptomyces coelicolor mutant for the cell division gene ftsZ. *Mol Microbiol* **14**: 243-254.

Merrick, M.J. (1976) A morphological and genetic mapping study of bald colony mutants of *Streptomyces coelicolor*. *J Gen Microbiol* **96**: 299-315.

Meyer, B.J., Maurer, R., and Ptashne, M. (1980) Gene regulation at the right operator (OR) of bacteriophage lambda. II. OR1, OR2, and OR3: their roles in mediating the effects.of repressor and cro. *J Mol Biol* **139**: 163-194.

Miguelez, E.M., Hardisson, C., and Manzanal, M.B. (1999) Hyphal death during colony development in *Streptomyces antibioticus*: morphological evidence for the existence of a process of cell deletion in a multicellular prokaryote. *J Cell Biol* **145**: 515-525.

Molle, V., and Buttner, M.J. (2000) Different alleles of the response regulator gene *bldM* arrest *Streptomyces coelicolor* development at distinct stages. *Mol Microbiol* **36**: 1265-1278.

Molle, V., Palframan, W.J., Findlay, K.C., and Buttner, M.J. (2000) WhiD and WhiB, homologous proteins required for different stages of sporulation in *Streptomyces coelicolor* A3(2). *J Bacteriol* **182**: 1286-1295.

Nguyen, K.T., Willey, J.M., Nguyen, L.D., Nguyen, L.T., Viollier, P.H., and Thompson, C.J. (2002) A central regulator of morphological differentiation in the multicellular bacterium *Streptomyces coelicolor*. *Mol Microbiol* **46**: 1223-1238.

Nguyen, K.T., Tenor, J., Stettler, H., Nguyen, L.T., Nguyen, L.D., and Thompson, C.J. (2003) Colonial differentiation in Streptomyces coelicolor depends on translation of a specific codon within the adpA gene. *J Bacteriol* **185**: 7291-7296.

Nodwell, J.R., McGovern, K., and Losick, R. (1996) An oligopeptide permease responsible for the import of an extracellular signal governing aerial mycelium formation in *Streptomyces coelicolor*. *Mol Microbiol* **22**: 881-893.

Nodwell, J.R., and Losick, R. (1998) Purification of an extracellular signaling molecule involved in production of aerial mycelium by *Streptomyces coelicolor*. *J Bacteriol* **180**: 1334-1337.

Nodwell, J.R., Yang, M., Kuo, D., and Losick, R. (1999) Extracellular complementation and the identification of additional genes involved in aerial mycelium formation in *Streptomyces coelicolor*. *Genetics* **151**: 569-584.

Noens, E.E., Mersinias, V., Traag, B.A., Smith, C.P., Koerten, H.K., and van Wezel, G.P. (2005) SsgA-like proteins determine the fate of peptidoglycan during sporulation of Streptomyces coelicolor. *Mol Microbiol* **58**: 929-944.

Nussbaum-Shochat, A., and Amster-Choder, O. (1999) BglG, the transcriptional antiterminator of the bgl system, interacts with the beta' subunit of the Escherichia coli RNA polymerase. *Proc Natl Acad Sci U S A* **96**: 4336-4341.

O'Connor, T.J., Kanellis, P., and Nodwell, J.R. (2002) The *ramC* gene is required for morphogenesis in *Streptomyces coelicolor* and expressed in a cell type-specific manner under the direct control of RamR. *Mol Microbiol* **45**: 45-57.

O'Connor, T.J., and Nodwell, J.R. (2005) Pivotal roles for the receiver domain in the mechanism of action of the response regulator RamR of Streptomyces coelicolor. *J Mol Biol* **351**: 1030-1047.

Ohnishi, K., Kutsukake, K., Suzuki, H., and Iino, T. (1990) Gene fliA encodes an alternative sigma factor specific for flagellar operons in Salmonella typhimurium. *Mol Gen Genet* **221**: 139-147.

Ohnishi, Y., Yamazaki, H., Kato, J.Y., Tomono, A., and Horinouchi, S. (2005) AdpA, a central transcriptional regulator in the A-factor regulatory cascade that leads to morphological development and secondary metabolism in Streptomyces griseus. *Biosci Biotechnol Biochem* **69**: 431-439.

Paget, M.S., Chamberlin, L., Atrih, A., Foster, S.J., and Buttner, M.J. (1999a) Evidence that the extracytoplasmic function sigma factor σ^E is required for normal cell wall structure in *Streptomyces coelicolor* A3(2). *J Bacteriol* **181**: 204-211.

Paget, M.S., Leibovitz, E., and Buttner, M.J. (1999b) A putative two-component signal transduction system regulates σ^E, a sigma factor required for normal cell wall integrity in *Streptomyces coelicolor* A3(2). *Mol Microbiol* **33**: 97-107.

Paget, M.S., Bae, J.B., Hahn, M.Y., Li, W., Kleanthous, C., Roe, J.H., and Buttner, M.J. (2001) Mutational analysis of RsrA, a zinc-binding anti-sigma factor with a thiol-disulphide redox switch. *Mol Microbiol* **39**: 1036-1047.

Plaskitt, K.A., and Chater, K.F. (1995) Influences of developmental genes on localized glycogen deposition in colonies of a mycelial prokaryote, *Streptomyces coelicolor* A3(2): a possible interface between metabolism and morphogenesis. *Phil Trans R Soc B* **347**: 105-121.

Pope, M.K., Green, B.D., and Westpheling, J. (1996) The *bld* mutants of *Streptomyces coelicolor* are defective in the regulation of carbon utilization, morphogenesis and cell-cell signalling. *Mol Microbiol* **19**: 747-756.

Pope, M.K., Green, B., and Westpheling, J. (1998) The *bldB* gene encodes a small protein required for morphogenesis, antibiotic production, and catabolite control in *Streptomyces coelicolor*. *J Bacteriol* **180**: 1556-1562.

Potuckova, L., Kelemen, G.H., Findlay, K.C., Lonetto, M.A., Buttner, M.J., and Kormanec, J. (1995) A new RNA polymerase sigma factor, σ^F, is required for the late stages of morphological differentiation in *Streptomyces* spp. *Mol Microbiol* **17**: 37-48.

Price, B., Adamidis, T., Kong, R., and Champness, W. (1999) A *Streptomyces coelicolor* antibiotic regulatory gene, *absB*, encodes an RNase III homolog. *J Bacteriol* **181**: 6142-6151.

Redenbach, M., Kieser, H.M., Denapaite, D., Eichner, A., Cullum, J., Kinashi, H., and Hopwood, D.A. (1996) A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3(2) chromosome. *Mol Microbiol* **21**: 77-96.

Ryding, N.J., Kelemen, G.H., Whatling, C.A., Flardh, K., Buttner, M.J., and Chater, K.F. (1998) A developmentally regulated gene encoding a repressor-like protein is essential for sporulation in *Streptomyces coelicolor* A3(2). *Mol Microbiol* **29**: 343-357.

Ryding, N.J., Bibb, M.J., Molle, V., Findlay, K.C., Chater, K.F., and Buttner, M.J. (1999) New sporulation loci in *Streptomyces coelicolor* A3(2). *J Bacteriol* **181**: 5419-5425.

Ryding, N.J., Anderson, T.B., and Champness, W.C. (2002) Regulation of the *Streptomyces coelicolor* calcium-dependent antibiotic by *absA*, encoding a cluster-linked two-component system. *J Bacteriol* **184**: 794-805.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular Cloning A Laboratory Manual.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

San Paolo, S., Huang, J., Cohen, S.N., and Thompson, C.J. (2006) rag genes: novel components of the RamR regulon that trigger morphological differentiation in Streptomyces coelicolor. *Mol Microbiol* **61**: 1167-1186.

Schmidt, R., Margolis, P., Duncan, L., Coppolecchia, R., Moran, C.P., Jr., and Losick, R. (1990) Control of developmental transcription factor sigma F by sporulation regulatory proteins SpoIIAA and SpoIIAB in Bacillus subtilis. *Proc Natl Acad Sci U S A* **87**: 9221-9225.

Schwedock, J., McCormick, J.R., Angert, E.R., Nodwell, J.R., and Losick, R. (1997) Assembly of the cell division protein FtsZ into ladder-like structures in the aerial hyphae of Streptomyces coelicolor. *Mol Microbiol* **25**: 847-858.

Shima, J., Penyige, A., and Ochi, K. (1996) Changes in patterns of ADP-ribosylated proteins during differentiation of Streptomyces coelicolor A3(2) and its development mutants. *J Bacteriol* **178**: 3785-3790.

Shin, D., and Groisman, E.A. (2005) Signal-dependent binding of the response regulators PhoP and PmrA to their target promoters in vivo. *J Biol Chem* **280**: 4089-4094.

Soliveri, J.A., Gomez, J., Bishai, W.R., and Chater, K.F. (2000) Multiple paralogous genes related to the *Streptomyces coelicolor* developmental regulatory gene *whiB* are present in *Streptomyces* and other actinomycetes. *Microbiology* **146**: 333-343.

Solomon, M.J., and Varshavsky, A. (1985) Formaldehyde-mediated DNA-protein crosslinking: a probe for in vivo chromatin structures. *Proc Natl Acad Sci U S A* **82**: 6470-6474.

Strauch, E., Takano, E., Baylis, H.A., and Bibb, M.J. (1991) The stringent response in Streptomyces coelicolor A3(2). *Mol Microbiol* **5**: 289-298.

Sun, J., Kelemen, G.H., Fernandez-Abalos, J.M., and Bibb, M.J. (1999) Green fluorescent protein as a reporter for spatial and temporal gene expression in *Streptomyces coelicolor* A3(2). *Microbiology* **145**: 2221-2227.

Takano, E., Tao, M., Long, F., Bibb, M.J., Wang, L., Li, W., Buttner, M.J., Deng, Z.X., and Chater, K.F. (2003) A rare leucine codon in adpA is implicated in the morphological defect of bldA mutants of Streptomyces coelicolor. *Mol Microbiol* **50**: 475-486.

Tan, H., and Chater, K.F. (1993) Two developmentally controlled promoters of *Streptomyces coelicolor* A3(2) that resemble the major class of motility-related promoters in other bacteria. *J Bacteriol* **175**: 933-940.

Tan, H., Yang, H., Tian, Y., Wu, W., Whatling, C.A., Chamberlin, L.C., Buttner, M.J., Nodwell, J., and Chater, K.F. (1998) The *Streptomyces coelicolor* sporulation-specific sigma WhiG form of RNA polymerase transcribes a gene encoding a ProX-like protein that is dispensable for sporulation. *Gene* **212**: 137-146.

Tan, H., Tian, Y., Yang, H., Liu, G., and Nie, L. (2002) A novel Streptomyces gene, samR, with different effects on differentiation of Streptomyces ansochromogenes and Streptomyces coelicolor. *Arch Microbiol* **177**: 274-278.

Tillotson, R.D., Wosten, H.A., Richter, M., and Willey, J.M. (1998) A surface active protein involved in aerial hyphae formation in the filamentous fungus *Schizophillum commune* restores the capacity of a bald mutant of the filamentous bacterium *Streptomyces coelicolor* to erect aerial structures. *Mol Microbiol* **30**: 595-602.

Traag, B.A., Kelemen, G.H., and Van Wezel, G.P. (2004) Transcription of the sporulation gene ssgA is activated by the IclR-type regulator SsgR in a whi-independent manner in Streptomyces coelicolor A3(2). *Mol Microbiol* **53**: 985-1000.

Tsao, S.W., Rudd, B.A., He, X.G., Chang, C.J., and Floss, H.G. (1985) Identification of a red pigment from *Streptomyces coelicolor* A3(2) as a mixture of prodigiosin derivatives. *J Antibiot* **38**: 128-131.

Ueda, K., Takano, H., Nishimoto, M., Inaba, H., and Beppu, T. (2005) Dual transcriptional control of amfTSBA, which regulates the onset of cellular differentiation in Streptomyces griseus. *J Bacteriol* **187**: 135-142.

van Wezel, G.P., van der Meulen, J., Kawamoto, S., Luiten, R.G., Koerten, H.K., and Kraal, B. (2000) *ssgA* is essential for sporulation of *Streptomyces coelicolor* A3(2) and affects hyphal development by stimulating septum formation. *J Bacteriol* **182**: 5653-5662.

van Wezel, G.P., and Vijgenboom, E. (2004) Novel aspects of signaling in Streptomyces development. *Adv Appl Microbiol* **56**: 65-88.

van Wezel, G.P., Krabben, P., Traag, B.A., Keijser, B.J., Kerste, R., Vijgenboom, E., Heijnen, J.J., and Kraal, B. (2006) Unlocking Streptomyces spp. for use as sustainable industrial production platforms by morphological engineering. *Appl Environ Microbiol* **72**: 5283-5288.

Vieira, J., and Messing, J. (1987) Production of single-stranded plasmid DNA. *Methods Enzymol* **153**: 3-11.

Viollier, P.H., Kelemen, G.H., Dale, G.E., Nguyen, K.T., Buttner, M.J., and Thompson, C.J. (2003) Specialized osmotic stress response systems involve multiple SigB-like sigma factors in *Streptomyces coelicolor. Mol Microbiol* **47**: 699-714.

White, J., and Bibb, M. (1997) *bldA* dependence of undecylprodigiosin production in *Streptomyces coelicolor* A3(2) involves a pathway-specific regulatory cascade. *J Bacteriol* **179**: 627-633.

Wietzorrek, A., and Bibb, M. (1997) A novel family of proteins that regulates antibiotic production in streptomycetes appears to contain an OmpR-like DNA-binding fold. *Mol Microbiol* **25**: 1181-1184.

Willey, J., Santamaria, R., Guijarro, J., Geistlich, M., and Losick, R. (1991) Extracellular complementation of a developmental mutation implicates a small sporulation protein in aerial mycelium formation by *S. coelicolor. Cell* **65**: 641-650.

Willey, J., Schwedock, J., and Losick, R. (1993) Multiple extracellular signals govern the production of a morphogenetic protein involved in aerial mycelium formation by *Streptomyces coelicolor. Genes Dev* **7**: 895-903.

Williams, J.G., and Mason, P.J. (1985) Hybridisation in the analysis of RNA. In *Nucleic acid hybridisation, a practical approach*. Hames, B.D. and S.J., H. (eds). Oxford: IRL Press, pp. 139-160.

Yamazaki, H., Ohnishi, Y., and Horinouchi, S. (2000) An A-factor-dependent extracytoplasmic function sigma factor (sigma(AdsA)) that is essential for morphological development in Streptomyces griseus. *J Bacteriol* **182**: 4596-4605.

Yamazaki, H., Ohnishi, Y., and Horinouchi, S. (2003) Transcriptional switch on of ssgA by A-factor, which is essential for spore septum formation in Streptomyces griseus. *J Bacteriol* **185**: 1273-1283.

Yanisch-Perron, C., Vieira, J., and Messing, J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**: 103-119.

Yudkin, M.D., and Clarkson, J. (2005) Differential gene expression in genetically identical sister cells: the initiation of sporulation in Bacillus subtilis. *Mol Microbiol* **56**: 578-589.

Zhen, L., and Swank, R.T. (1993) A simple and high yield method for recovering DNA from agarose gels. *Biotechniques* **14**: 894-898.

Appendices

Appendix A: Potential BldD targets identified by microarray analysis of RNA isolated from the *S. coelicolor bldD* mutant (1169) after induced *bldD* overexpression

A number of potential BldD targets were identified by the microarray analysis. Although only a subset were chosen for transcript, DNA binding, and mutational analyses, two genes later found to be of significance are discussed here. The original microarray data are presented in Figure 3.8 and Table A.1, and the plotted data are found in Figure A.1. Descriptions of the genes are found in Table A.2.

A.1 *sco0751* (*scf81.10c*) as a potential BldD target

sco0751 (*scf81.10c*) encodes a putative regulator with a predicted phosphatase domain homologous to the phosphatase domains of RsbU and SpoIIE of *B. subtilis* (http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml). RsbU is one of the phosphatases that dephosphorylates the anti-sigma factor antagonist RsbV, activating it and indirectly activating the master stress-response sigma factor σ^B (Kuo *et al.*, 2004). Likewise, SpoIIE dephosphorylates the anti-sigma factor SpoIIAA, freeing it to interact with the anti-sigma factor SpoIIAB and release the sporulation-specific sigma factor σ^B (Yudkin and Clarkson, 2005). Transcription of *sco0751* showed a similar pattern as that of the known *bldD*

Table A.1 BldD-induced changes in gene expression of known and putative _bldD_ target genes

Gene name*	Cosmid ID	SCO number	pAU244						pIJ6902				
			t=15	t=30	t=45	t=60	t=120		t=15	t=30	t=45	t=60	t=120
argB	SCL24.14c	SCO1578	-0.3	-0.09	-1.05	-1.61			2.4	-0.2	0.36	0.21	0.38
argC	n/a	SCO1580	0.95	-0.62		-1.57	-1.28		0.82	0.62	-0.16	0.69	-1.1
argD	SCL24.13c	SCO1577	-0.28	-0.12	0.2	-1.91	-1.23		1.48	-0.18		0.28	-1.44
argG	SC4G1.02	SCO7036	-0.1	-0.76		-2.08	-1.28		1.64	0.83	0.59	0.88	-0.49
argH	SCL24.06c	SCO1570	0.04	0.13	-0.38	-1.12	-0.36		2.15	0.54	0.1	0.23	-0.04
argJ	SCL24.15c	SCO1579	0.49	-0.99	-0.86	-0.9	-0.5		1.7	0.05	0.01	-0.33	0.22
argR	SCL24.12c	SCO1576	0.62	0.12		-0.96	-0.1		0.07	0.43	-0.13	-0.16	-0.5
bdtA	SCE68.26c	SCO3328	0.1	-0.21	-0.17	0.05	0.01		0.28	0.1	-0.04	-0.13	0.13
bldD	SC9C5.13	SCO1489	0.31	1.42		1.8	1.65		0.77	0.39	0.45	-0.01	0.48
bldN	SCE68.21	SCO3323	0.33	0.7	-0.21	-0.84	-1.33		0.91	1.52	0.29	0.39	0.31
*_cseB_	SCE94.09	SCO3358	-0.35	0.22	-0.16	1.34	0.46		-0.35	-0.57	-0.33	-0.7	-0.3
*_cseC_	SCE94.10	SCO3359	-0.27	0.04	-0.07	1.66	0.45		-0.1	-0.21	-0.45	-0.37	0.23
leuC	n/a	SCO5553	-0.46	-0.66	-0.71	-2.4	-1.36		1.68	1.09	0.54	0.09	-0.08
*_paaI_	SCBAC17A6.03c	SCO7470	-0.06	-0.28	-0.15	-0.35	-0.14		-0.03	-0.16	-0.17	-0.13	-0.85
paaK	SCBAC17A6.02c	SCO7469	0.23	-0.54	-0.26	-0.82	-1.05		-0.4	-0.29	-0.08	0.57	-0.69
rplR	SCD31.43	SCO4718	-0.15	-0.41	0.38	-1.65	-1.01		0.24	-0.3	0.15	-0.04	-0.03
sigE	SCE94.07	SCO3356	0.31	1.15	0.41	2.92	1.43		-1.15	-0.27	0.34	-0.26	0.59
sigH	2SC7G11.05c	SCO5243	0.04	-0.13	-1.00	-1.13	-1.98		-0.10	0.11	0.00	0.5	-0.01
whiG	SC2E1.38	SCO5621	-0.21	0.27	-0.72	-0.30			0.08	0.25	-0.47	-0.39	
	2SC7G11.24	SCO5262	0.12	0.93	0.83	1.53	1.12		0.09	0.39	-0.21	-0.16	0.29
*	2SC7G11.31	SCO5269	-0.26		-0.38	-0.86	-0.06		-0.31	-0.85	-0.82	-0.47	0.2
*	2SC7G11.32	SCO5270	-0.78	-0.69		-0.84	-0.77		0.11	-1.08	-1.36	-1.01	-1.13
*	2SC7G11.34	SCO5272	0.59	0.37	-0.92	-1	-0.66		0.06	0.64	0.61	0.05	0.86
	2SC7G11.35	SCO5273	0.3	0.34	-0.6	-0.69	-1.36		0.11	0.77	0.04	0.32	0.58
*spot 1	SC2C3.01/ SC1H10.23	SCO7034	-0.27		0.03	-0.59	-0.56		-1.01		0.38	0.09	-2.17
*spot 2	SC2C3.01/ SC1H10.23	SCO7034	-0.11	-0.49		-0.9	-0.13		-0.47	-2.01	-0.14	-1	-0.22

Table A.1 continued

Gene name*	Cosmid ID	SCO number	pAU244						pIJ6902				
			t=15	t=30	t=45	t=60	t=120		t=15	t=30	t=45	t=60	t=120
	SC3C3.04	SCO5718	-0.23	-0.23	-0.55	-0.67	-1.16		0.4	0.68	0.53	-0.29	0.03
	SC4H2.17	SCO5796	-0.58	0.4	0.83	1.3	1.3		-0.2	-2.18	0.15	-0.24	0.5
	SC4H2.18c	SCO5797	0.2	0.9	0.45	2.46	0.11		-0.28	0.05	0.01	-0.68	-0.29
	SC4H2.19c	SCO5798	0.58	1.5	0.91	3.3	0.64		0.34	0.49	0.5	-0.11	-0.52
	SCC117.02	SCO2529	-0.2	-0.07	-1.31	-2.02	-2.6		-0.23	-0.1	-0.9	-0.38	-0.26
	SCD12A.23	SCO4340	-0.32	-0.67	-1.23	-0.74	-2.17		-0.35	-0.77	-0.54	-0.96	0.65
	SCD19.03c	SCO4348	-0.66	0.24	0.32	-0.96	-2.57		-0.22	-0.66	-0.34	-0.88	-0.41
	SCD19.04c	SCO4349	-0.43	1.24	0.3	-0.63	-2.33		-2.33	0.13	0.18	0.01	-1.08
	SCD8A.12c	SCO4239	-0.8	0.55	-0.53	-2.35	-2.79		-0.46	-0.22	-0.5	-0.92	-1.18
	SCE65.05c	SCO3469	-0.12	0.01	-1.47	-1.66	-2.38		-0.08	0.12	-0.68	-0.69	-0.44
*	SCE68.25c	SCO3327	0.7	0.38	-0.01	-1.27	0.41		0.64	0.24	-0.51	0.46	
	SCF81.10c	SCO0751	0.41	0.25	-0.59	-1.19	-0.56		-0.04	0.01	0.34	0.78	-0.31
	SCG8A.17c	SCO1163	-0.07	0.78	-0.34	-0.66	-1.17		-0.42	-1.36	-0.2	-0.06	-0.46
*spot1	SCH35.15	SCO3709	-0.73	0.02	-0.58		-0.12		-0.75	-1.97	0.26	-1.11	-1.54
*spot2	SCH35.15	SCO3709	0.35	-0.46	0.26	-0.88	-0.52		-0.61	-0.24	-0.27	0.57	-0.83
	SCH35.16	SCO3708	0.29	-0.36	-0.33	-1.44	-1.47		0.02	-0.17	-1.31	0.23	-1.99
	SCH5.23	SCO3560	0	-0.08	-0.72	-1.77	-2.91		0.59	0.01	0.5	-0.14	0.6
	SCI5.07	SCO1799	0.49	0.5	-0.12	-1.32	-1.99		-0.07	0.39	-0.33	-0.46	0.06
	SCJ12.07	SCO0195	1.03	0.64	1.26	0.78	-1.52		0.48		-2.3	-2.71	-3.08
*	SCL24.05c	SCO1569	-0.28	0.56					0.11	0.2		0.3	
	SCL6.31	SCO1474	-0.07	1.02	-0.26	-1.06	-1.65		0.56	0.26	0.13	-0.54	0.2
	SCM10.10c	SCO0922	0.7	1.21	-1.3	2.99	2.12		-0.33	-1.13	-1.22	-1.17	0.75
	SCM10.12c	SCO0924		0.09	-3.08	1.56	0.68		-0.92	-2.47	-2.01	-2.48	-0.17

* Genes examined because of their position relative to those identified as targets by initial analysis of microarray data. No data were available for 2SC7G11.33, SCD19.05c or SCM10.11c. Empty cells indicate unavailable data.

Numbers represent the $\log_2$ fold change in expression of genes in the indicated samples relative to the uninduced sample. Positive numbers indicate induction while negative numbers indicate repression of transcription.

Table A.2 Description of known and putative BldD target genes

Gene name*	Cosmid ID	SCO number	Function**	BldD binding sequence Consensus match
argB	SCL24.14c	SCO1578	Acetylglutamate kinase	none (operon with *argC*)
argC	n/a	SCO1580	N-acetyl-gamma-glutamyl-phosphate reductase	CGTGC (4) GCACG
**argD*	SCL24.13c	SCO1577	Acetonitrile aminotransferase	none (operon with *argC*)
argG	SC4G1.02	SCO7036	Arginosuccinate synthase	ACTCA ACACC
argH	SCL24.06c	SCO1570	Argininosuccinate lyase	GGTGC (15) GCACC
argJ	SCL24.15c	SCO1579	Putative glutamate N-acetyltransferase	none (operon with *argC*)
**argR*	SCL24.12c	SCO1576	Arginine repressor	none (operon with *argC*)
bdtA	SCE68.26c	SCO3328	Putative DNA-binding protein	AGTGA GCACG (13) TCACC
bldD	SC9C5.13	SCO1489	DNA binding protein	AGTAA (12) TCACA
bldN	SCE68.21	SCO3323	ECF sigma factor	AGTGC (2) GCACG, GGTGA (13) GCACG
**cseA*	SCE94.08	SCO3357	Hypothetical protein	Operon with *sigE*
**cseB*	SCE94.09	SCO3358	Two-component system response regulator	Putative operon with *sigE*
**cseC*	SCE94.10	SCO3359	Putative sensory histidine kinase	Putative operon with *sigE*
**gabD*	SC2C3.02/SC4G1.01	SCO7035	Succinate-semialdehyde dehydrogenase	none - putative operon with *argG*
leuC	n/a	SCO5553	3-isopropylmalate dehydratase large subunit	none
**paaI*	SCBAC17A6.03c	SCO7470	Putative phenylacetic acid degradation protein PaaI	GGTGT (6) CCACT
paaK	SCBAC17A6.02c	SCO7469	Phenylacetate-CoA ligase	none (putative operon with *paaI*)
rplR	SCD31.43	SCO4718	50S ribosomal protein L18	Putative operon with SCD31.40
sigE	SCE94.07	SCO3356	ECF sigma factor	AGTGG (11) CCACC GGTGC (21) GCACC
sigH	2SC7G11.05c	SCO5243	Sigma factor	AGTGA (2) GCACG
whiG	SC2E1.38	SCO5621	Sigma factor	AGTGA TCACC, AGTCA (8) TCACG
	2SC7G11.24	SCO5262	Putative dehydrogenase	GGTGT (3) ACACG
*	2SC7G11.30	SCO5268	Hypothetical protein	TGTGA (8) GCACA

Table A.2 continued

Gene name*	Cosmid ID	SCO number	Function**	BldD binding sequence Consensus match
*	2SC7G11.31	SCO5269	Hypothetical protein	None- operon with 2SC7G11.30
*	2SC7G11.32	SCO5270	Conserved hypothetical protein	CGTGG (12) CCACC
*	2SC7G11.33	SCO5271	Conserved hypothetical protein	GGTGA (24) CCACG
*	2SC7G11.34	SCO5272	Hypothetical protein	none - putative operon with 2SC7G11.30 or .33
	2SC7G11.35	SCO5273	Hypothetical protein	none - putative operon with 2SC7G11.30 or .33
*	SC2C3.01/SC1H10.23	SCO7034	Putative aminotransferase	none - putative operon with *argG*
	SC3C3.04	SCO5718	Putative signalling molecule	CGTGA TCACC
	SC4H2.17	SCO5796	Conserved hypothetical protein (induced)	GGTGG (3) GCACG (4) ACACG
	SC4H2.18c	SCO5797	Putative secreted protein (induced)	None - putative operon with SC4H2.19c
	SC4H2.19c	SCO5798	Putative secreted protein (induced)	AGTGT (16) TCACC AGTCA ACACC
	SCC117.02	SCO2529	Putative metalloprotease	AGTGA (8) TCACT
	SCD12A.23	SCO4340	Putative integrase	GGTGT (9) CCACA 32 bp into the coding region
	SCD19.03c	SCO4348	Hypothetical protein	None (putative operon with SCD19.04c or .05c)
	SCD19.04c	SCO4349	Conserved hypothetical protein	AGTGA (2) TCACC
*	SCD19.05c	SCO4350	Putative integrase	AGTGA (23) TCACT
	SCD8A.12c	SCO4239	Putative small membrane protein	CGTGA (1) GCACT CCACC
	SCE65.05c	SCO3469	Transposase, contains TTA	ACTGT (18) TCACA
*	SCE68.25c	SCO3327	Hypothetical protein	Operon with *bdtA*
	SCF81.10c	SCO0751	Putative regulatory protein	GGTGC (7) CCACC
	SCG8A.17c	SCO1163	Hypothetical protein	ACTCT (11) TCACT
*	SCH35.15	SCO3709	Putative ECF sigma factor	TGTGG (17) GCACC
	SCH35.16	SCO3708	Putative integral membrane protein	None (putative operon with SCH35.15)
	SCH5.23	SCO3560	Putative ATP-binding protein	AGTGG (6) TCACC
	SCI5.07	SCO1799	Hypothetical protein	none

Table A.2 continued

Gene name*	Cosmid ID	SCO number	Function**	BldD binding sequence Consensus match
	SCJ12.07	SCO0195	Putative lipoprotein Note: appears induced	GGTCG (11) GCACA GGTGA (18) CCACC
*	SCL24.05c	SCO1569	Putative oxidoreductase	CGTGG (7) GCACG) within the *argH* coding region
	SCL6.31	SCO1474	Hypothetical protein	GGTCA (11) ACACA
	SCM10.10c	SCO0922	Putative reductase iron-sulphur protein	GGTGA (5) CCACC (2) GCACA
*	SCM10.11c	SCO0923	putative reductase flavoprotein subunit	none
	SCM10.12c	SCO0924	putative cytochrome B subunit	CGTGG CCACC AGTGC (16) CCACG

* Genes examined because of their position relative to those identified as targets by initial analysis of the microarray data.

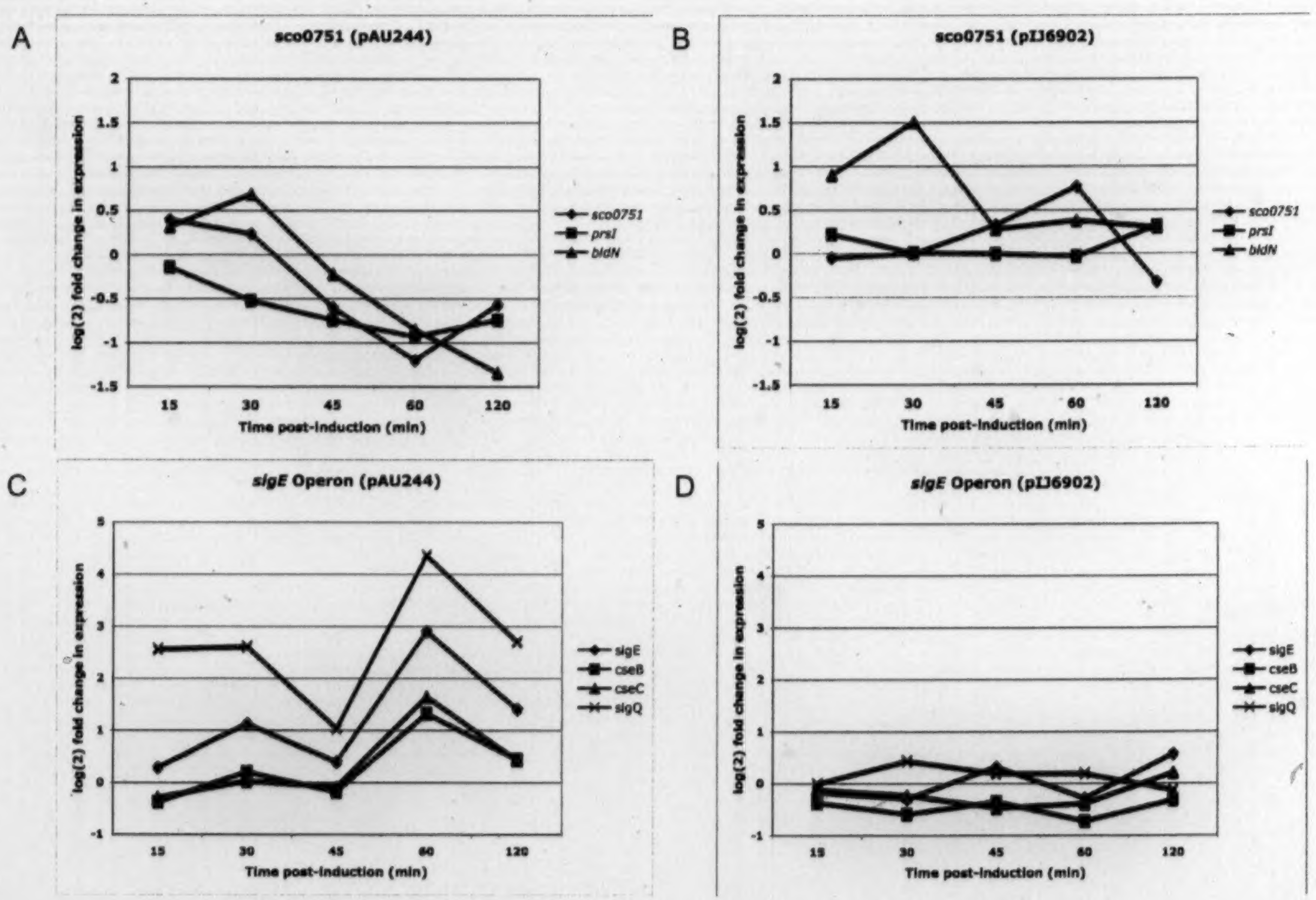
** ECF: extracytoplasmic function

Figure A.1

Microarray results (see Table A.1), plotted as the $\log_2$-fold change in expression relative to the uninduced sample at each time point post-induction. Values above zero indicate induction and values below zero indicate repression. The plots on the left side of the page (A and C) represent the samples harvested from the *bldD* mutant strain containing inducible *bldD* (1169 with pAU244), while the plots on the right side of the page (B and D) represent the samples harvested from the *bldD* mutant strain containing the control plasmid (1169 with pIJ6902).

A) and B) The putative BldD target *sco0751*, with *prsI* and *bldN* included for comparison.

C) and D) The *sigE* operon (includes *sigE*, *cseA*, *cseB* and *cseC*, but no data were available for *cseA*), with *sigQ* included for comparison.


A
sco0751 (pAU244)
log(2) fold change in expression
2
1.5
1
0.5
0
-0.5
-1
-1.5
15
30
45
60
120
Time post-induction (min)
sco0751
prsI
bldN
B
sco0751 (pIJ6902)
log(2) fold change in expression
2
1.5
1
0.5
0
-0.5
-1
-1.5
15
30
45
60
120
Time post-induction (min)
sco0751
prsI
bldN
C
sigE Operon (pAU244)
log(2) fold change in expression
5
4
3
2
1
0
-1
15
30
45
60
120
Time post-induction (min)
sigE
cseB
cseC
sigQ
D
sigE Operon (pIJ6902)
log(2) fold change in expression
5
4
3
2
1
0
-1
15
30
45
60
120
Time post-induction (min)
sigE
cseB
cseC
sigQ

target gene, *bldN* (Fig. A.1 A). As with *bldN*, *sco0751* transcription did not change over time in the control strain (Fig. A.1 B). Regulation by BldD is likely direct, as there is a putative BldD binding site 119 bp upstream of the translation start codon of *sco0751* (GGTGC [7] CCACC) (Table A.2). Discussion of the potential role of the product of *sco0751* in regulating the activity of ArsI can be found in section 4.1.

A.2 *sigE* (*sco3356, sce94.07*) as a potential BldD target

One of the surprising genes identified as a potential BldD target by the microarray analysis was *sigE*, which encodes an extracytoplasmic function sigma factor required for integrity of the cell wall (Hong *et al.*, 2002; Paget *et al.*, 1999b). Positive regulators of *sigE* have been identified, so further induction of *sigE* expression by BldD seems redundant. *sigE* is known to be transcribed as a monocistronic transcript and also as an operon with at least three downstream genes: *cseA* (*sce94.08/sco3357*), *cseB* (*sce94.09/sco3358*) and *cseC* (*sce94.10/sco3359*) (Paget *et al.*, 1999a; Paget *et al.*, 1999b). All three downstream genes were therefore included in the analysis (Tables A.1 and A.2). Looking at all of the genes in an operon provides a level of internal control, especially important when finding an unexpected or unprecedented result; if the result seen is real, then all of the genes of an operon should have the same response. *sigE* was clearly induced following *bldD* induction, following the same transcriptional pattern as *sigQ*, though only reaching eight-fold induction in contrast to the 21-fold induction seen with *sigQ* (Fig. A.1, C and D). Consistent with the operon structure, the pattern of induction seen in the downstream genes *cseB* and *cseC* was the same (Fig. A.1, C and D), although no data were available for *cseA*. Induction of the *sigE* operon appeared to be significant, and given the presence of two potential BldD binding sites upstream of *sigE*, the cause is likely direct BldD regulation. The first putative binding site (AGTGG [11] CCACC) begins 48 bp downstream of the transcription start point, and the second putative BldD binding site (GGTGC [21] GCACC) starts eight base pairs within the coding region. Further exploration of the significance of BldD regulation of *sigE* can be found in the discussion (section 4.4).

Appendix B: Putative targets of BldD chosen for PCR-targeted mutagenesis

Two genes were chosen for mutational analysis in addition to the BldD target gene disruptions described in the body of this thesis. Both had putative BldD binding sites upstream of the start codon, and both were repressed during *bldD* induction, as measured by microarray analysis. The original image from the microarray analysis is found in Figure 3.8, and Figure B.1 contains plots illustrating the data. Raw data are found in Table A.1.

B.1 *sco5718* (*sc3c3.04*) as a potential BldD target

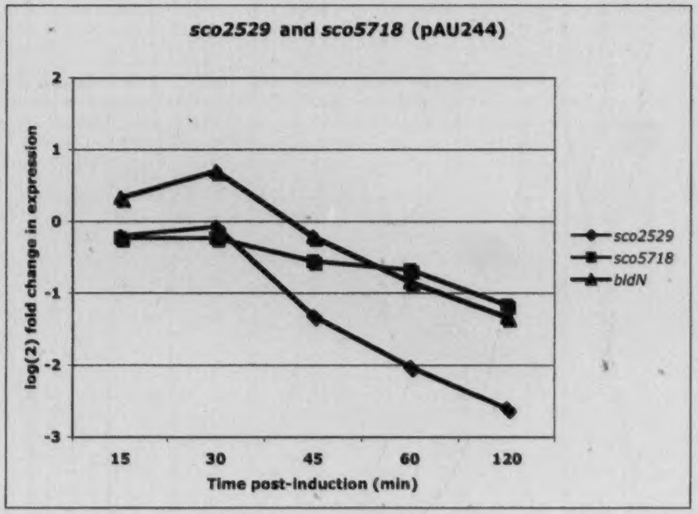
sco5718 (*sc3c3.04*) encodes a putative secreted protein. Secreted signalling molecules play an important role in *S. coelicolor* development, so it seemed reasonable that this gene might be a BldD target. There was a good BldD binding consensus sequence upstream of the start codon - CGTGA TCACC. The transcription start point has not been mapped, so the position of the putative binding site within the promoter cannot be discussed.

Although no transcript mapping was carried out, *sco5718* was deleted by PCR-targeted mutagenesis (Gust *et al.*, 2003). This work was carried out by L. Bui as a summer student under the direction of C. Galibois. The apramycin resistance cassette was amplified using the primers CGA45 and CGA46 (Table 2.4). Replacement of the wild type gene in the cosmid with the disrupted gene was confirmed by restriction digest. The mutant cosmid was then transferred by conjugation into the wild type *S. coelicolor* strain M600 as described above

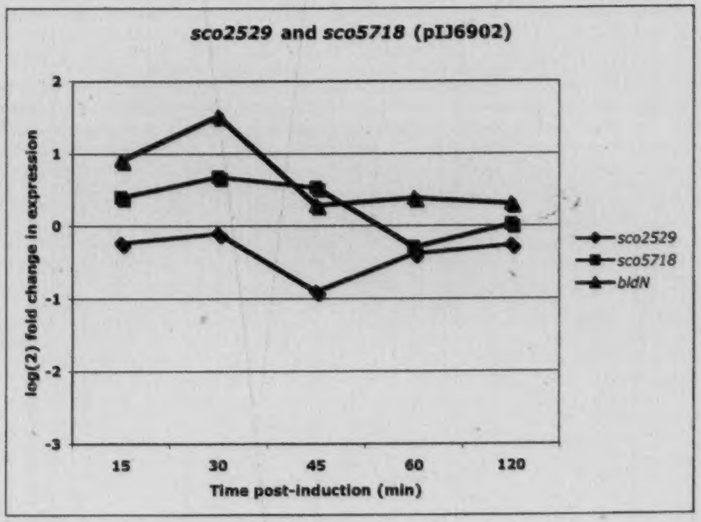
Figure B.1

sco2529 and *sco5718* are repressed in response to *bldD* induction. Microarray results for *sco2529* and *sco5718* (see Table A.1) are plotted as the $\log_2$-fold change in expression relative to the uninduced sample at each time point post-induction. Values above zero indicate induction and values below zero indicate repression. Plot A represents the samples harvested from the *bldD* mutant strain containing inducible *bldD* (1169 with pAU244), while plot B represents the samples harvested from the *bldD* mutant strain containing the control plasmid (1169 with pIJ6902). The known BldD target *bldN* is included for comparison.

A


sco2529 and *sco5718* (pAU244)
log(2) fold change in expression
2
1
0
-1
-2
-3
15
30
45
60
120
Time post-induction (min)
sco2529
sco5718
bldN

B


sco2529 and *sco5718* (pIJ6902)
log(2) fold change in expression
2
1
0
-1
-2
-3
15
30
45
60
120
Time post-induction (min)
sco2529
sco5718
bldN

(Section 3.6). Thirty transconjugants were screened, and 20 apramycin esistant, kanamycin sensitive colonies were isolated. All apramycin resistant, kanamycin sensitive colonies were expected to represent colonies in which a double crossover recombination had occurred to replace the wild type gene in the chromosome with the disrupted copy. Since all exhibited the same visual phenotype, one of the twenty colonies was selected for further analysis, and after three generations a spore stock was made to generate colonies that arose from single genomes. Chromosomal DNA was isolated from nine such kanamycin-sensitive colonies, and from one kanamycin-resistant colony to serve as a control. Complete replacement of the wild type gene in the resulting strains was confirmed by separate Southern analyses of *Bam*HI-digested chromosomal DNA. A gene specific probe (1779 bp *Bam*HI fragment from cosmid 3C3) allowed visualization of the difference in mobility between the wild type and disrupted copies of the gene, the apramycin cassette was used to confirm the presence of the disrupted copy, and the wild type cosmid was used to check for gross rearrangements (Fig. B.2). Strain 18-1 was chosen for phenotypic analysis. The null mutant displayed a wild type phenotype on a variety of media, shown in Figure B.3 by growth on R2YE. When the strain was subjected to osmotic shock, the mutant strain showed the same growth defects as the wild type parent strain M600 (Fig. B.4). Because the phenotype gave no indication of the role of the gene, in development or elsewhere, no further analysis was pursued.

Figure B.2

Confirmation of replacement of the wild type *sco5718/sc3c3.04* gene with the apramycin resistance cassette in the chromosome of *S. coelicolor*. This work was performed by L. Bui under the direction of C. Galibois. Chromosomal DNA was isolated from mutants and digested with *Bam*HI before separation by electrophoresis on an agarose gel, transfer to a nylon membrane, and hybridization with a radioactively-labelled probe (see below). The marker used in all cases was lambda phage DNA digested with *Pst*I (λPst).

A) The gene-specific probe used was a 1779 bp *Bam*HI restriction fragment from cosmid SC3C3. Shown are the hybridizing wild-type (smaller) and disrupted (larger) fragments resulting from the restriction digest of the chromosomal DNA.

B) The apramycin cassette was used as a probe to confirm the presence of the cassette in the double (kanamycin sensitive) and single (kanamycin resistant) crossover mutants. Similarly digested chromosomal DNA from the wild type *S. coelicolor* (M600) and the undisrupted cosmid (SC3C3) were included as negative controls. The disrupted cosmid (Δ*sco5718*) was included as a positive control.

C) Wild type cosmid (SC3C3) was used as a probe to confirm the absence of any gross chromosomal rearrangements caused by the recombination that resulted in replacement of *sco5718* with the apramycin resistance disruption cassette..

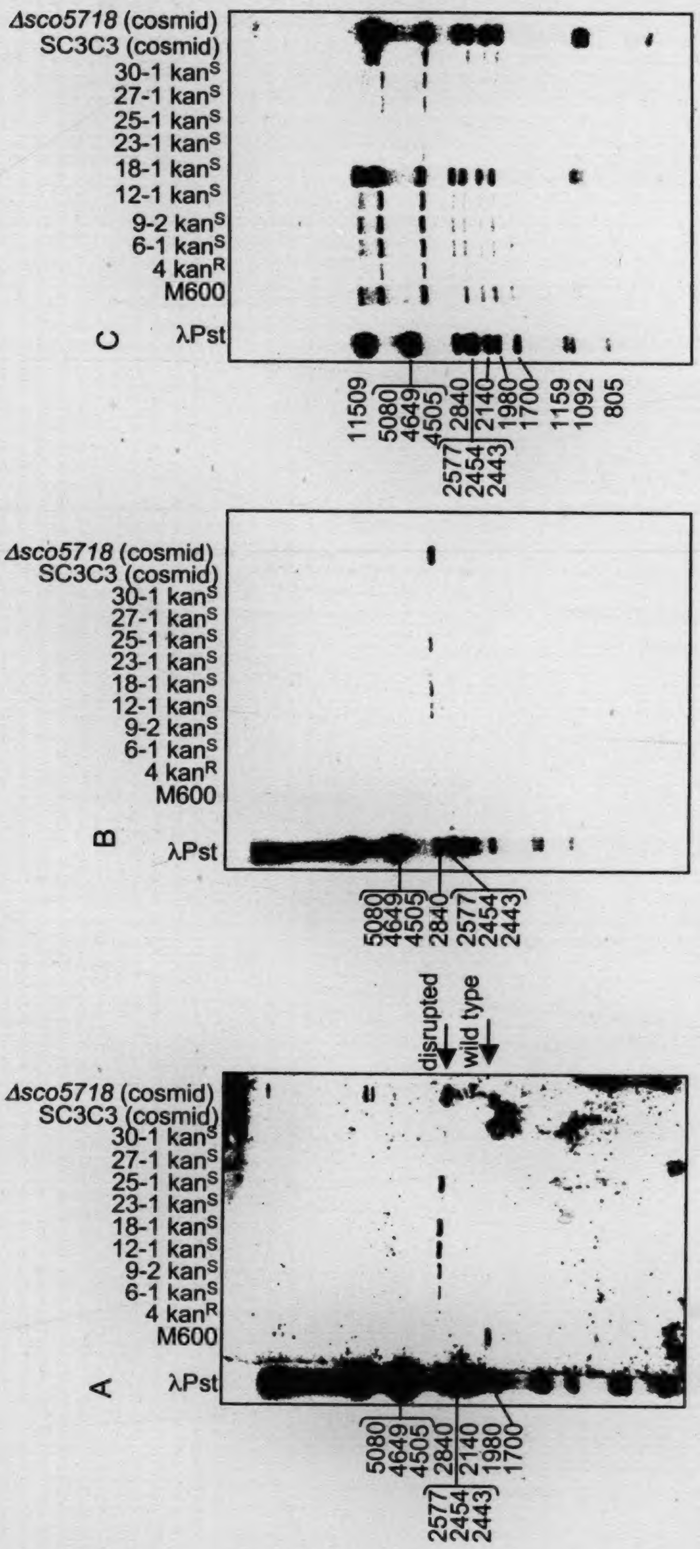
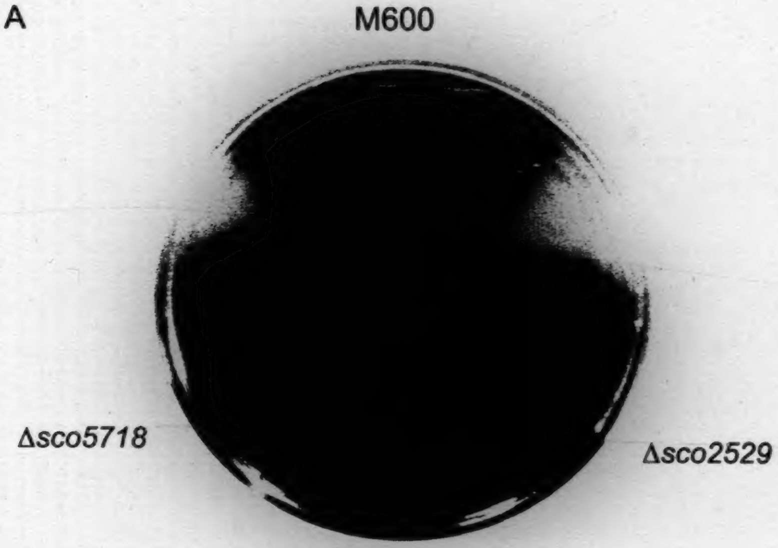

C
Δsco5718 (cosmid)
SC3C3 (cosmid)
30-1 kan^S
27-1 kan^S
25-1 kan^S
23-1 kan^S
18-1 kan^S
12-1 kan^S
9-2 kan^S
6-1 kan^S
4 kan^R
M600
λPst
11509
5080
4649
4505
2840
2577
2454
2443
2140
1980
1700
1159
1092
805
B
Δsco5718 (cosmid)
SC3C3 (cosmid)
30-1 kan^S
27-1 kan^S
25-1 kan^S
23-1 kan^S
18-1 kan^S
12-1 kan^S
9-2 kan^S
6-1 kan^S
4 kan^R
M600
λPst
5080
4649
4505
2840
2577
2454
2443
disrupted
wild type
A
Δsco5718 (cosmid)
SC3C3 (cosmid)
30-1 kan^S
27-1 kan^S
25-1 kan^S
23-1 kan^S
18-1 kan^S
12-1 kan^S
9-2 kan^S
6-1 kan^S
4 kan^R
M600
λPst
5080
4649
4505
2840
2577
2454
2443
2140
1980
1700

Figure B.3

Null mutants of *sco5718* and *sco2529* do not show developmental or antibiotic defects. Null mutants of *sco5718* and *sco2529*, with the congenic wild type parent strain M600. Strains were grown on R2YE for three days. Plates are shown in the bottom (A) and top (B) views. This work was performed by L. Bui under the direction of C. Galibois.


A
M600
Δsco5718
Δsco2529

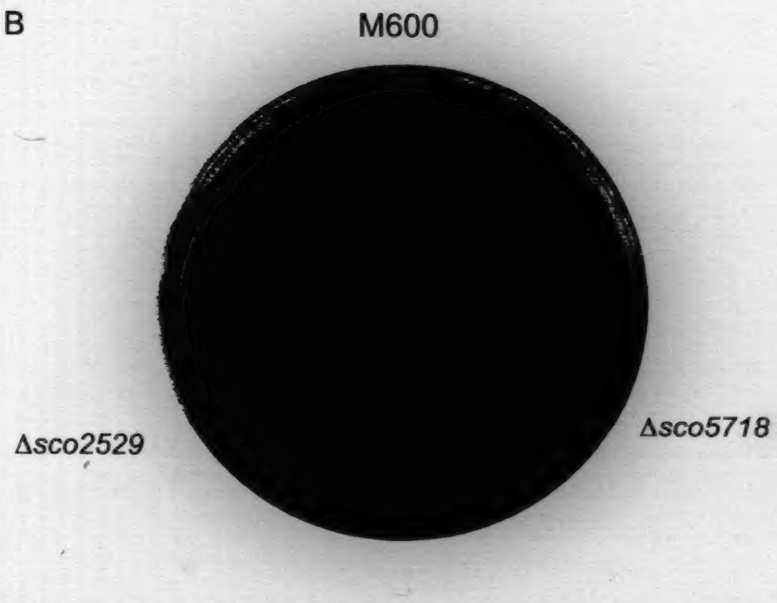

B
M600
Δsco2529
Δsco5718

Figure B.4

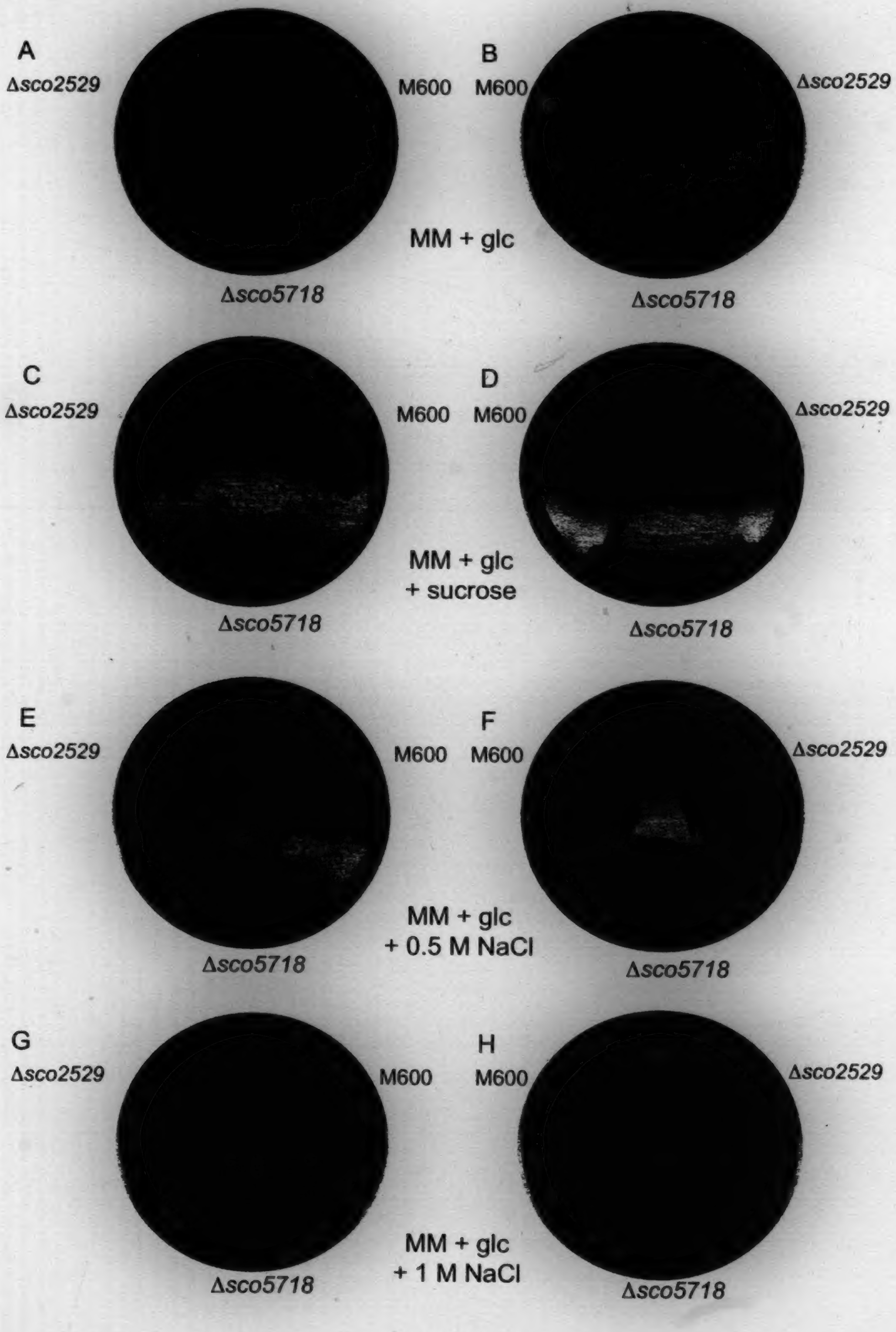
Δ*sco2529* and Δ*sco5718* mutants show no change in response to osmotic shock. Osmotic shock testing of was performed on the Δ*sco2529* and Δ*sco5718* mutants and compared to the parent strain M600. Strains were grown for 48 h on minimal medium with glucose (MM + glc) or the same medium supplemented with either NaCl or sucrose.

A and B) Front and back views, respectively, grown on MM + glc.

C and D) Front and back views, respectively, grown on MM + glc supplemented with 1 M sucrose (MM + glc + sucrose).

E and F) Front and back views, respectively, grown on MM + glc supplemented with 0.5 M NaCl (MM + glc + 0.5 M NaCl).

G and H) Front and back views, respectively, grown on MM + glc supplemented with 1 M NaCl (MM + glc + 1 M NaCl).


A
Δsco2529
M600
Δsco5718
B
M600
Δsco2529
Δsco5718
MM + glc
C
Δsco2529
M600
Δsco5718
D
M600
Δsco2529
Δsco5718
MM + glc
+ sucrose
E
Δsco2529
M600
Δsco5718
F
M600
Δsco2529
Δsco5718
MM + glc
+ 0.5 M NaCl
G
Δsco2529
M600
Δsco5718
H
M600
Δsco2529
Δsco5718
MM + glc
+ 1 M NaCl

B.2 *sco2529* (*scc117.02*) as a potential BldD target

sco2529 (*scc117.02*) encodes a putative metalloprotease. Since there was a very good BldD binding consensus upstream of the start codon [AGTCA (6) TCACG], it was likely a direct BldD target. This gene was chosen for analysis because of a potentially important role in development. The extracytoplasmic function sigma factor σ^{BldN} is very unusual in that it is synthesized as a proprotein with an 86 residue N-terminal extension which is cleaved by an unknown protease to release the active sigma factor (Bibb and Buttner, 2003). In previously reported work, two candidate metalloproteases had been disrupted to check for a bald (*bldN* disruption) phenotype. Disruptions of both appeared wild type (Bibb and Buttner, 2003), leaving open the question of which protease is responsible for σ^{BldN} processing. Given that *bldN* transcription is directly regulated by BldD, it might seem redundant for BldD to also regulate the protease responsible for σ^{BldN} cleavage, but regulation in *S. coelicolor* is rife with redundancies and it still seemed reasonable to determiine if the product of *sco2529* might play that role.

L. Bui, under the supervision of C. Galibois, deleted *sco2529* using the method of Gust *et al.* (2003). She used CGA56 and CGA57 to amplify the apramycin cassette, and confirmed replacement of the wild type gene with the disruption cassette in the cosmid by digestion with *Eco*RI and *Sac*I (Fig. 3.25). Conjugation was again used to move the disrupted cosmid from *E. coli* ET12567/pUZ8002 into *S. coelicolor* strain M600. Sixteen exconjugants were grown as above (section 3.8.2), and the one kanamycin-sensitive colony was selected and grown for two subsequent generations. Spores were collected from this strain, plated to extinction for single colonies, and chromosomal DNA isolated from three of the resulting single colonies as well as from one kanamycin-resistant isolate. Southern analyses were performed on *KpnI/PstI* digested chromosomal DNA, confirming that the wild type gene had been replaced with the apramycin cassette in strains 3-2-1 and 6, and that there were no gross chromosomal rearrangements in the vicinity of *sco2529* (Fig. B.5). The gene-specific probe was the 1829 bp *KpnI/XhoI* fragment from cosmid C117, and the apramycin cassette and the cosmid itself were used as probes. Strain 3-2-1 was chosen for phenotypic analysis. The null mutants displayed a wild type phenotype on a variety of media, including R2YE (Fig. B.3), and did not look any different than the wild type parent (M600) when subjected to osmotic shock (Fig. B.4). Had *sco2529* encoded the metalloprotease responsible for cleavage of *bldN*, the expected phenotype of the null mutant would have been bald on all of the media tested; clearly, this was not the case. No further analysis was pursued for this target.

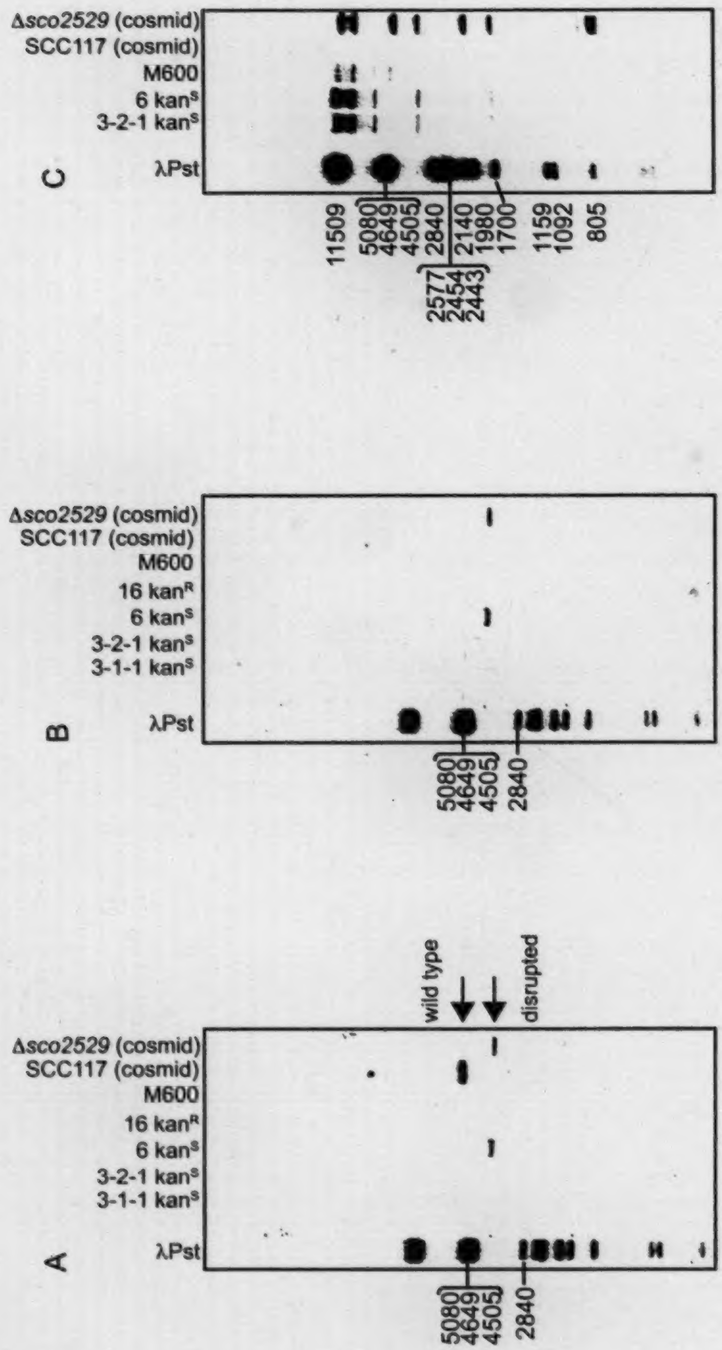
Figure B.5

Confirmation of replacement of the wild type *sco2529/scc117.02* gene with the apramycin resistance cassette in the chromosome of *S. coelicolor*. This work was performed by L. Bui under the direction of C. Galibois. Chromosomal DNA was isolated from mutants and digested with *Kpn*I and *Pst*I before separation by electrophoresis on an agarose gel, transfer to a nylon membrane, and hybridization with a radioactively-labelled probe. The marker used in all cases was lambda phage DNA digested with *Pst*I (λPst).

A) The gene-specific probe used was the 1829 bp *Kpn*I/*Xho*I fragment from cosmid SCC117. Shown are the hybridizing wild-type (larger) and disrupted (smaller) fragments resulting from the restriction digest of the chromosomal DNA.

B) The apramycin cassette was used as a probe to confirm the presence of the cassette in the double (kanamycin sensitive) and single (kanamycin resistant) crossover mutants. Similarly digested chromosomal DNA from the wild type *S. coelicolor* (M600) and the undisrupted cosmid (SCC117) were included as negative controls. The disrupted cosmid (Δ*sco2529*) was included as a positive control.

C) Wild type cosmid (SCC117) was used as a probe to confirm the absence of any gross chromosomal rearrangements caused by the recombination that resulted in replacement of *arsI* with the apramycin resistance disruption cassette.


C
Δsco2529 (cosmid)
SCC117 (cosmid)
M600
6 kan^S
3-2-1 kan^S
λPst
11509
5080
4649
4505
2840
2577
2454
2443
2140
1980
1700
1159
1092
805
B
Δsco2529 (cosmid)
SCC117 (cosmid)
M600
16 kan^R
6 kan^S
3-2-1 kan^S
3-1-1 kan^S
λPst
5080
4649
4505
2840
A
wild type
disrupted
Δsco2529 (cosmid)
SCC117 (cosmid)
M600
16 kan^R
6 kan^S
3-2-1 kan^S
3-1-1 kan^S
λPst
5080
4649
4505
2840

Appendix C: *prsI* overexpression attempt

Because of the difficulty in generating a deletion of the wild-type *prsI* gene, an attempt was made to generate a strain of *S. coelicolor* with a second copy of *prsI* under the control of an inducible promoter. Such a strain would be useful to generate a strain where the chromosomal copy of *prsI* could be deleted, in the presence of the induced copy, then the induction removed and the null phenotype observed. Such an experiment would also serve to demonstrate whether or not *prsI* is an essential gene.

The vector chosen for construction of the inducible *prsI* overexpression plasmid was pIJ6902, the same vector used to overexpress *bldD* in the microarray experiments. This vector contains a thiostrepton-inducible promoter, p*tipA*, ahead of the polylinker. To introduce a gene for overexpression, the start codon must be maintained. To this end, the upstream primer (CGA64) was designed to contain an *Nde*I site (which contains an in-frame ATG sequence that can act as a start codon) followed by 19 bp homologous to the *prsI* coding sequence immediately downstream of the native *prsI* start codon. The downstream primer (CGA65) was designed to amplify the gene from 60 bp downstream of the *prsI* stop codon. The location of both primers is illustrated in Figure 3.11. Using these primers, a 583 bp fragment was amplified by PCR from an 879 bp fragment of cosmid SCE25 produced by *Eco*RI/*Not*I digestion, using Expand Long Template polymerase (Roche). Expand Long Template polymerase was used despite the potential for errors in amplification, because the higher fidelity enzyme (Expand High Fidelity polymerase, Roche) was unable to amplify a product using this primer set. The resulting 583 bp PCR product was cloned into pCR®2.1 TOPO (Invitrogen), generating pAU247, and confirmed to be error-free by sequencing.

pAU247 was digested with *Eco*RI and the *prsI* fragment ligated into *Eco*RI-digested pIJ2925 to generate pAU249 and pAU252 (containing the fragment in opposite orientations). From this point, the work done was carried out by K. Gislason, as a technician under the direction of C. Galibois. pAU249 was digested with *Nde*I and *Bgl*II and the *prsI* fragment introduced into pIJ6902 to produce pAU259. This final construct was passed through the non-methylating *E. coli* strain ET12567 (containing pUZ8002) before being introduced by conjugation into the wild type *S. coelicolor* strain M145.

The resulting *S. coelicolor* strain was plated on R2YE with 30 μg/mL thiostrepton (the same concentration used for the microarray induction). Neither the vector control (M145 with pIJ6902) nor the test strain (M145 with pAU259) differentiated properly, or produced wild type levels of antibiotics. This meant that no conclusions could be drawn from the experiment; any phenotype observed could have been a result of growth on thiostrepton rather than due to the expression of the inducible copy of *prsI*.

Appendix D: BldD overexpression/purification attempts

The protein used for all of the experiments described in the main body of the thesis was His_6-tagged BldD purified by Marie Elliot, a previous graduate student, in 1999 (Elliot, 2000). There was some concern, especially with the later EMSAs and footprinting experiments, that the age of the protein was giving misleading results and indicating weaker binding than would be seen with fresher protein. With that in mind, many attempts were made to purify fresh protein. These attempts were carried out at various times by C. Galibois, K. Gislason and T. Locke. The methods described previously were used (Elliot and Leskiw, 1999; Elliot, 2000). In short, *E. coli* JM109 containing pQE9BldD$^+$ [pQE9 (Qiagen) containing the *bldD* coding sequence] was grown overnight to saturation, then diluted 1/100 into fresh LB with ampicillin to an OD_{600} of 0.6-0.7 and induced by the addition of 0.5 mM IPTG and grown for an additional two hours. Cells were harvested by centrifugation at 4°C, then washed and resuspended in five volumes of sonication buffer [50 mM sodium phosphate (pH7.8), 300 mM NaCl] with 1x Complete EDTA-free protease inhibitor cocktail (Roche). Cells were either used fresh or stored in 3 mL aliquots at -20°C. When needed, aliquots were thawed on ice and sonicated 3 x 20 sec, full duty cycle, setting 2-3 (Branson Sonifier, VWR Scientific) or lysed with one passage through a French Press (American Instrument Company, Silver Spring, Maryland). Extracts were then clarified by centrifugation and 10 mM imidazole added before purification by Ni-affinity chromatography. Extracts were mixed with Ni-NTA resin (Qiagen) equilibrated with wash buffer (50 mM NaH_2PO_4, pH8.0, 300 mM NaCl, 20 mM imidazole) and incubated at 4°C prior loading in a disposable BioRad column. The column was washed twice with wash buffer, first unmodified then with imidazole added to 50 mM. The protein was then eluted with elution buffer (50 mM NaH_2PO_4, pH8.0, 300 mM NaCl) containing 100 mM imidazole (first two elutions) or 350 mM imidazole (last two elutions). Aliquots from all steps were denatured in protein loading dye and electrophoresed on a 15% gel to determine the purity of the protein with representative results shown in Fig. D.1 A. Samples containing a significant amount of pure BldD were concentrated by spin column (Millipore Amicon Ultra, 10 kDa cutoff) and either used immediately or stored at 4°C for a maximum of four days before use.

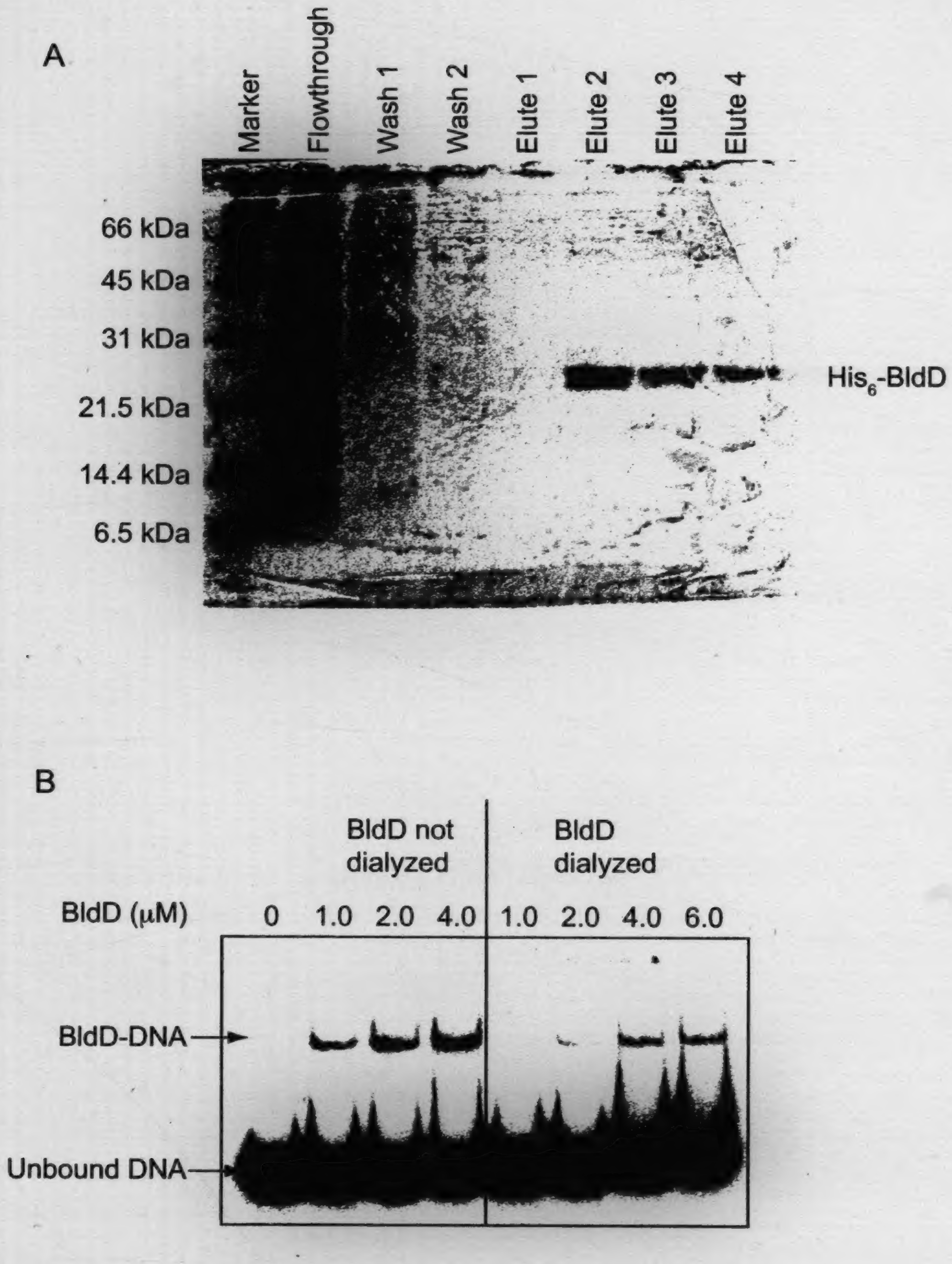
EMSAs were performed to test the activity of the protein before the longer, more complicated DNase I footprinting experiments were attempted. Standard conditions were used, as described in Section 2.6.1. The probe for the *arsI* promoter (CGA47-48) was used because it had given good results in earlier work. Although some amount of shifted DNA can be seen, more than half of the probe remains unshifted even with 4-6 μM BldD. One sample of BldD was dialyzed overnight against 20 mM Tris-HCl (pH8.5), 150 mM NaCl and 10% glycerol (fresh protease inhibitor added to the dialyzed sample) in case the imidazole was inhibiting the activity of the protein prior to use for EMSA. The dialyzed protein had even less binding activity than the undialyzed protein, probably due to the overnight refrigeration and dilution of the protease inhibitor involved in the dialysis. Many more attempts were made, using numerous variations on this basic protocol, but the results shown in Fig. D.1 B are typical of

Figure D.1

BldD purified by nickel affinity chromatography. .

A) Aliquots taken from nickel-column purification of His$_6$-BldD, denatured and analyzed by SDS-PAGE (15% gel). Flowthrough is cellular protein with no affinity to the column, Wash 1 contains proteins removed from the column by 20 mM imidazole and Wash 2 contained 50 mM imidazole. BldD was eluted with two volumes of elution buffer containing 100 mM imidazole(Qiagen) and two volumes of elution buffer containing 350 mM imidazole. Band size was determined using the BioRad Broad Range SDS-Page marker.

B) EMSA performed with freshly purified BldD. The first lane contains only end-labelled probe (CGA47-48, containing the BldD binding site upstream of *arsI*). The next three lanes contain probe bound by up to 4 μM freshly purified, undialyzed BldD, and the final four lanes contain probe bound by 1-6 μM freshly purified BldD, dialyzed to remove imidazole. Bound and unbound DNA are indicated by arrows.


A
Marker
Flowthrough
Wash 1
Wash 2
Elute 1
Elute 2
Elute 3
Elute 4
66 kDa
45 kDa
31 kDa
21.5 kDa
14.4 kDa
6.5 kDa
His_6-BldD
B
BldD not dialyzed
BldD dialyzed
BldD (μM)
0
1.0
2.0
4.0
1.0
2.0
4.0
6.0
BldD-DNA
Unbound DNA

all of the experiments. The original construct was also sequenced to determine if the inactivity of the protein was due to mutations accumulated over storage and culturing. No sequence changes were seen.

The results from these attempts to purify fresh BldD were very disappointing, and meant that it was not possible to repeat EMSA experiments (such as CGA49-42 - longer *prsI* probe) or to perform DNase I footprinting on *sigQ*.